

**Theoretische und empirische Abhandlungen über die Wahrnehmung und
Wirkung der Ertragsbesteuerung sowie über die Nutzung ausgewählter
steuerlicher Begünstigungsnormen**

Inaugural-Dissertation

zur Erlangung des akademischen Grades eines Doktors der Wirtschaftswissenschaft des
Fachbereichs Wirtschaftswissenschaft der Freien Universität Berlin

vorgelegt von

Diplom-Ökonom Christian Sielaff

aus Recklinghausen, wohnhaft in Berlin

2011

Tag der Disputation: 02.12.2011

Erstgutachter: Univ.-Prof. Dr. Jochen Hundsdoerfer
Institut für Betriebswirtschaftliche Prüfungs- und Steuerlehre
Professur für Betriebswirtschaftslehre
insbesondere Betriebswirtschaftliche Steuerlehre
Freie Universität Berlin
Garystr. 21
14195 Berlin

Zweitgutachter: Univ.-Prof. Dr. Harald Jansen
Lehrstuhl für ABWL, Betriebswirtschaftliche Steuerlehre /
Wirtschaftsprüfung
Friedrich-Schiller-Universität Jena
Carl-Zeiss-Straße 3
07743 Jena

Vorbemerkungen

Die vorliegende Dissertationsschrift wurde nach der Promotionsordnung des Fachbereichs Wirtschaftswissenschaft vom 16.07.2008 angefertigt. Die Dissertationsschrift wurde kumulativ nach der Ausführungsvorschrift für das kumulative Promotionsverfahren, welche der Promotionsausschuss des Fachbereichs Wirtschaftswissenschaft am 09.07.2008 verabschiedet hat, angefertigt.

Hierzu erkläre ich Folgendes:

1. Zu Beginn der Dissertationsschrift erfolgt eine zusammenhängende Darstellung (Exposé) der gesamten Thematik gemäß § 9 Abs. 2 Buchstabe b der Promotionsordnung vom 16.07.2008. Die zusammenhängende Darstellung enthält eine Einführung in die Problemstellung, erläutert die wesentlichen Forschungsfragen und fasst die grundlegenden Ergebnisse der Einzelbeiträge zusammen.
 2. In die Einleitung wurde gemäß § 9 Abs. 2 Buchstabe b der Promotionsordnung vom 16.07.2008 eine tabellarische Übersicht aufgenommen, die den Anteil der Eigenleistung bei Artikeln, die in Ko-Autorenschaft entstanden sind, kenntlich macht.
 3. Eingereichte und bereits veröffentlichte Fachartikel sind der Dissertation als Kopie beigelegt. Für noch nicht veröffentlichte Beiträge enthält die Dissertation die entsprechende Arbeitspapier- bzw. Manuskriptfassung.
 4. Die eingereichten Fachartikel wurden entsprechend den Anforderungen gemäß Nr. 4 Buchstaben a bis c der Ausführungsvorschrift für das kumulative Promotionsverfahren vom 09.07.2008 ausgewählt.
 5. Ich versichere, dass die Dissertation von mir selbstständig erstellt und der Eigenanteil bei Artikeln in Ko-Autorenschaft wahrheitsgemäß angegeben wurde. Die Arbeit hat keiner anderen Prüfungsbehörde vorgelegen.
-

Exposé

Inhaltsverzeichnis

Inhaltsverzeichnis.....	II
Tabellenverzeichnis.....	III
1. Einleitung	1
2. Thematische Einordnung der Forschungsbeiträge anhand der verschiedenen Forschungsziele.....	5
3. Forschungsfragen der Einzelbeiträge	9
3.1 Die gefühlte Steuerbelastung des Einkommens – Eine empirische Analyse für Erwerbstätige in Deutschland.....	9
3.2 The Influence of Tax Labeling and Tax Earmarking on the Willingness to Contribute – A Conjoint Analysis	10
3.3 Steuerkomplexität und Arbeitsangebot - Eine experimentelle Untersuchung	12
3.4 Progressionseffekte und Varianten zur optimalen Steuerplanung bei der Thesaurierungsbegünstigung – Eine Abweichungsanalyse	13
3.5 Charitable Giving in the German Welfare State: Fiscal Incentives and Crowding Out.....	14
Literaturverzeichnis.....	IV
Zusammenfassung.....	X
Summary	XII
Lebenslauf.....	XIII
Publikationsverzeichnis.....	XIV
Überblick über die Einzelbeiträge.....	XV

Tabellenverzeichnis

Tabelle 1 - Übersicht über die Einzelbeiträge	1
Tabelle 2 - Einordnung und Abgrenzung der Einzelbeiträge.....	2

1. Einleitung

Die vorliegende Dissertationsschrift besteht aus fünf Einzelbeiträgen, welche sich mit ausgewählten Fragen der Betriebswirtschaftlichen Steuerlehre beschäftigen. Tabelle 1 enthält eine Übersicht über die Einzelbeiträge dieser Dissertationsschrift sowie Angaben über Ko-Autorenschaften und Informationen zur Veröffentlichung.

Titel	Ko-Autoren	Eigener Anteil	Zeitschrift, Fundstelle
Die gefühlte Steuerbelastung des Einkommens – Eine empirische Analyse für Erwerbstätige in Deutschland*	Blaufus, Kay Bob, Jonathan Hundsdoerfer, Jochen Kiesewetter, Dirk Weimann, Joachim	25%	Steuer und Wirtschaft 2010 (4), S. 337-345
The Influence of Tax Labeling and Tax Earmarking on the Willingness to Contribute - A Conjoint Analysis*	Blaufus, Kay Hundsdoerfer, Jochen Kiesewetter, Dirk Weimann, Joachim	33%	arqus Diskussionsbeitrag, August 2011
Steuerkomplexität und Arbeitsangebot - Eine experimentelle Untersuchung	---	100%	Diskussionsbeitrag des Fachbereichs Wirtschaftswissenschaft (FACTS), Freie Universität Berlin, Nr. 2011/13
Progressionseffekte und Varianten zur optimalen Steuerplanung bei der Thesaurierungsbegünstigung – Eine Abweichungsanalyse	Hechtner, Frank Hundsdoerfer, Jochen	33%	Schmalenbachs Zeitschrift für betriebswirtschaftliche Forschung 2011 (5), S. 214-239
Charitable Giving in the German Welfare State: Fiscal Incentives and Crowding Out	Bönke, Timm Massarrat-Mashhadi, Nima	33%	Public Choice 2011, (online first)

Tabelle 1 - Übersicht über die Einzelbeiträge

* Diese Arbeiten sind Teil eines größeren Gemeinschaftsprojektes mehrerer Lehrstühle für das Bundesministerium der Finanzen. Die tatsächlichen Anteile der einzelnen Autoren an diesen beiden Einzelbeiträgen weichen deshalb vom Kehrwert der Anzahl aller auf diesem Artikel vermerkten Autoren deutlich ab.

Tabelle 2 gibt einen Überblick über die Einordnung und Abgrenzung der fünf Einzelbeiträge dieser Dissertationsschrift anhand thematischer und methodischer Kriterien.

Beitrag	Thema	Methode	Untersuchungsdesign	Untersuchungsgegenstand
Die gefühlte Steuerbelastung des Einkommens – Eine empirische Analyse für Erwerbstätige in Deutschland	Steuerwahrnehmung	Empirie	Befragung	Steuerliche Belastungswirkung
The Influence of Tax Labeling and Tax Earmarking on the Willingness to Contribute - A Conjoint Analysis			Conjoint-Analyse	Steuerliche Belastungswirkung
Steuerkomplexität und Arbeitsangebot - Eine experimentelle Untersuchung			Experiment	Steuerliche Belastungswirkung
Progressionseffekte und Varianten zur optimalen Steuerplanung bei der Thesaurierungsbegünstigung – Eine Abweichungsanalyse	Entscheidungswirkungen der Besteuerung	Theorie	Modelltheorie	Steuerliche Begünstigungsnorm
Charitable Giving in the German Welfare State: Fiscal Incentives and Crowding Out		Empirie	Datenanalyse	Steuerliche Begünstigungsnorm

Tabelle 2 - Einordnung und Abgrenzung der Einzelbeiträge

Die Dissertationsschrift beschäftigt sich mit der Wahrnehmung der Besteuerung bei den Zensiten sowie dem Einfluss der Besteuerung auf ökonomische Entscheidungen. Während die klassische wirtschaftswissenschaftliche Literatur häufig auf der Annahme aufbaut, dass Wirtschaftssubjekte die Besteuerung vollständig verstehen und somit Steuern in rationaler Weise in ihren Entscheidungs- und Planungskalkülen berücksichtigen, untersucht die jüngere

Forschung immer mehr die subjektiv wahrgenommene Steuerlast und ihre Wirkungen.¹ Die aus theoretischen Entscheidungsmodellen abgeleiteten Ergebnisse werden dabei auf ihre Gültigkeit in der Realität untersucht, um anschließend die zugrunde liegenden Modelle über deren Implikationen entweder empirisch zu bestätigen oder aber die Modelle an die Realität anzupassen. An diesen Forschungsstrang knüpft die vorliegende Dissertationsschrift an. Zunächst wird empirisch untersucht, wie das deutsche Einkommensteuersystem und das davon ausgehende Belastungsgefühl von den Steuerpflichtigen wahrgenommen wird, bevor in dem zweiten Beitrag Möglichkeiten aufgezeigt werden, wie dieses Belastungsgefühl durch gezielte Maßnahmen beeinflusst werden kann. Im Speziellen wird untersucht, ob sich die individuelle Zahlungsbereitschaft bei der Steuererhebung durch die Benennung und die Zweckbindung einer Steuer beeinflussen lässt. Der dritte Beitrag untersucht im Rahmen eines Experiments den Einfluss steuerlicher Tarifkomplexität auf die Arbeits-Freizeit-Entscheidung von Steuerpflichtigen. Damit ergänzt dieser Beitrag die ersten beiden Beiträge zur Steuerwahrnehmung durch die experimentelle Messung der aus der subjektiven Steuerwahrnehmung resultierenden Entscheidungswirkungen. Dieser Beitrag geht somit über die reine Beobachtung der individuellen Steuerwahrnehmung hinaus und lässt sich folglich, wie in Tabelle 2 ersichtlich, thematisch auf der Schnittstelle zwischen „Steuerwahrnehmung“ und „Entscheidungswirkung der Besteuerung“ ansiedeln.

Die folgenden zwei Beiträge (Beitrag 4 und 5) widmen sich dann der Analyse ausgewählter steuerlicher Begünstigungsnormen. Nachdem zunächst die Frage im Mittelpunkt stand, wie Steuern wahrgenommen werden und welche Auswirkungen die Komplexität des Steuersystems auf das Entscheidungsverhalten hat, werden im Weiteren gezielt steuerliche Begünstigungsnormen herausgegriffen und untersucht. Dabei wird zum einen anhand einer Begünstigungsnorm theoretisch gezeigt, wie komplex eine vollständige Steuerplanung in der Realität sein kann. Zum anderen wird anhand einer weiteren Begünstigungsnorm empirisch untersucht, inwieweit diese in der Realität die gewünschte Wirkung entfaltet und somit fiskalpolitisch gerechtfertigt werden kann.

¹ So unterstellt bspw. bereits *Ramsey* (1927), dass Steuerpflichtige auf Steueränderungen in der gleichen Weise reagieren wie auf Preisänderungen. Diese Annahme wird auch in zahlreichen weiteren Beiträgen innerhalb der wirtschaftswissenschaftlichen Forschung übernommen, bspw. *Harberger* (1964), *Mirrlees* (1971), *Atkinson/Stiglitz* (1976). Arbeiten zur tatsächlichen Wahrnehmung der Besteuerung finden sich bereits bei *Schmölders* (1960) oder in der jüngeren Literatur bei *Hundsdoerfer* und *Sichtmann* (2007), *Blaufus* und *Ortlieb* (2009), *Cogden et al.* (2009). Eine ausführliche Darstellung zum Verhalten von Steuerpflichtigen bietet außerdem *Kirchler* (2007).

Der vierte Beitrag analysiert die optimale Nutzung der Thesaurierungsbegünstigung des § 34a EStG und verdeutlicht an diesem Beispiel, wie komplex eine Steuerplanung unter besonderer Berücksichtigung von Progressionseffekten sein kann. Es handelt sich bei diesem Beitrag um eine modelltheoretische Analyse, die exemplarisch am Beispiel der Thesaurierungsbegünstigung aufzeigt, welcher Planungsaufwand notwendig ist, um eine rationale Steuerplanung durchzuführen. Im Ergebnis stützt dieser Beitrag die untersuchten Hypothesen der ersten drei Einzelbeiträge, indem gezeigt wird, wie Steuerpflichtige handeln müssten, wenn diese den rationalen Annahmen klassischer Steuerplanungsmodelle entsprechen wollten. Diese Komplexität lässt sich als eine der Ursachen für die Abweichung zwischen der tatsächlichen und der subjektiv wahrgenommenen Steuerbelastung interpretieren, welche die Kernforschungsfrage der ersten drei Beiträge darstellt. Dadurch knüpft dieser Beitrag an die zuvor dargestellten Beiträge thematisch an.

Der fünfte Beitrag widmet sich ebenfalls einer steuerlichen Begünstigungsnorm. Es wird anhand einer empirischen Analyse am Beispiel der Spendenabzugsbegünstigung des § 10b EStG untersucht, welche fiskalpolitischen Auswirkungen durch diese Steuerbegünstigung entstehen und inwieweit sich diese Begünstigungsnorm aus fiskalpolitischer Sicht rechtfertigen lässt. Die Spendenabzugsfähigkeit stellt eine steuerliche Begünstigungsnorm dar, welche von Steuerpflichtigen in Anspruch genommen werden kann und diese motivieren soll, einen eigenen Beitrag zur Finanzierung besonders förderungswürdiger und somit steuerbegünstigter Zwecke zu leisten.² Wenn sich diese Vorschrift allerdings aus fiskalpolitischer Sicht nicht rechtfertigen lässt, sollte darauf ggf. aus Vereinfachungsgründen im deutschen Steuerrecht verzichtet werden.

Diese Untersuchung liegt dabei auf der Schnittmenge von Betriebswirtschaftlicher Steuerlehre und der Finanzwissenschaft, wobei der Schwerpunkt meines Erachtens eher bei der Finanzwissenschaft angesiedelt ist.³

Im folgenden Kapitel wird diese Dissertationsschrift thematisch in die bisherige Forschung eingebettet, bevor im Anschluss daran die Einzelbeiträge hinsichtlich Forschungsfrage, Methodik und der daraus gewonnen Erkenntnisse genauer dargestellt werden.

² Vgl. §§ 52 - 54 AO.

³ Eine umfassendere thematische Einordnung der Beiträge in die einzelnen Forschungsdisziplinen findet sich im folgenden Kapitel.

2. Thematische Einordnung der Forschungsbeiträge anhand der verschiedenen Forschungsziele

Um die einzelnen Forschungsbeiträge thematisch einzuordnen, muss zunächst dargestellt werden, wie sich die Aufgaben und Ziele der Betriebswirtschaftlichen Steuerlehre⁴ voneinander abgrenzen lassen. Eine neuere Abgrenzung findet sich bei *Hundsdoerfer/Kiesewetter/Sureth* (2008).⁵ Diese unterteilen die Forschungsziele der Betriebswirtschaftlichen Steuerlehre in vier Bereiche.⁶ Zu diesen Bereichen zählen die Steuerbelastungsmessung⁷, die modellgestützte Steuerplanungslehre⁸, die empirische Steuerwirkungslehre⁹ und die Steuerrechtsgestaltungslehre¹⁰. Insbesondere die empirische Steuerwirkungslehre stellt dabei einen bisher in der Betriebswirtschaftlichen Steuerlehre nur wenig erforschten Bereich dar.¹¹ Dieser beschäftigt sich mit der empirischen Überprüfung von Steuerwirkungen und versucht dabei Informationen über die tatsächlichen Verhaltensanpassungen von Steuerpflichtigen als Reaktion auf die Besteuerung zu gewinnen. Dieses Wissen ist notwendig, um die aus theoretischen Modellen abgeleiteten Ergebnisse entweder zu bestätigen oder aber die Modelle zu verbessern und an das tatsächliche Verhalten anzupassen, wenn die aus den Modellen abgeleiteten Hypothesen über das Verhalten der Steuerpflichtigen in der Realität im Rahmen der empirischen Überprüfung falsifiziert werden können.¹²

In der Literatur wird im Rahmen ökonomischer Modelle vorrangig davon ausgegangen, dass Steuerpflichtige ihre Steuerbelastung exakt berechnen können und ausgehend von diesen Berechnungen ihre ökonomischen Entscheidungen treffen.¹³ Allerdings ist in der jüngeren

⁴ Die Betriebswirtschaftliche Steuerlehre selbst wird dabei i.d.R. als Teildisziplin der Allgemeinen Betriebswirtschaftslehre gesehen. Vgl. *Jacobs* (2004), S. 251.

⁵ Kritisch dazu *Schmiel* (2009).

⁶ Eine ähnliche Abgrenzung findet sich auch bei *Fischer/Scheeloch/Sigloch* (1980).

⁷ Bei der Steuerbelastungsmessung lässt sich weiterhin eine rechtliche und eine wirtschaftliche Steuerbelastung unterscheiden. Vgl. *Schneider* (2002), S. 10f.

⁸ *Hundsdoerfer/Kiesewetter/Sureth* (2008), S. 63 verstehen darunter das Ableiten von Handlungsempfehlungen an Unternehmer. Inwieweit unter den Unternehmerbegriff auch einzelne Individuen zu subsumieren sind vgl. *Schneider* (1995), S. 30f. sowie *Hechtner* (2010), S. 14.

⁹ Aufgabe der Steuerrechtsgestaltungslehre ist es, gegenüber dem Gesetzgeber Empfehlungen zur Ausgestaltung des Steuerrechts auszusprechen. Vgl. *Hundsdoerfer/Kiesewetter/Sureth* (2008), S. 65. Zu dieser Aufgabe auch *Wagner/Schwenk* (2003), S. 374-376.

¹⁰ Vgl. *Wagner* (1974), S. 396.

¹¹ Vgl. *Wagner* (2004), S. 243f.

¹² Methodisch kann sich bei diesem Vorgehen die Theoriebildung bzw. Überprüfung durch Induktion und Deduktion vermischen. Vgl. zur Experimentalmethodik *Westermann* (2000), S. 31f.

¹³ Vgl. Fußnote 1.

Literatur im Bereich der Betriebswirtschaftlichen Steuerlehre eine klare Trendwende hin zur empirischen Überprüfung solcher Modelle erkennbar.¹⁴ Dies führt dazu, dass die klassischen Modellannahmen häufig überdacht werden müssen.¹⁵ Dabei verwischen zunehmend auch die Grenzen zwischen der Betriebswirtschaftlichen Steuerlehre und ihren Nachbardisziplinen hinsichtlich Methodik und Fragestellung.¹⁶ Die ersten drei Einzelbeiträge dieser Dissertationsschrift erweitern den Bereich der empirischen Steuerwirkungslehre und leisten einen neuen Forschungsbeitrag, indem Hypothesen über das Verhalten von Steuerpflichtigen und die Wahrnehmung des Steuersystems aufgestellt und empirisch überprüft werden. Dabei steht als zentrale Frage die subjektive Wahrnehmung und die daraus resultierenden Wirkungen der Besteuerung bei den Steuerpflichtigen im Mittelpunkt. Methodisch wird dabei auf das Instrument der Befragung (Einzelbeitrag 1), der Conjoint-Analyse (Einzelbeitrag 2) sowie des Experiments (Einzelbeitrag 3) zurückgegriffen. Damit weist dieser Teil der vorliegenden Dissertationsschrift ein breit aufgestelltes Spektrum bei der Datengewinnung und Auswertung auf. Sämtliche genutzte Daten in diesem Themenkomplex wurden selbst erhoben.

Die genutzten Methoden werden dabei zum Teil aus Nachbardisziplinen entnommen, was ein Zeichen für die Interdisziplinarität bei der noch relativ jungen empirischen Steuerwirkungslehre ist.¹⁷ Die Nutzung solcher Methoden für Fragestellungen der Betriebswirtschaftlichen Steuerlehre leistet somit an sich bereits einen eigenen Beitrag zur Weiterentwicklung der Betriebswirtschaftlichen Steuerforschung.

Die zwei weiteren Einzelbeiträge beschäftigen sich mit der Nutzung steuerlicher Begünstigungsnormen. Der erste Beitrag, der sich mit der optimalen Nutzung der Thesaurierungsbegünstigung des § 34a EStG unter Beachtung von Progressionseffekten beschäftigt, gehört zu dem Bereich der Steuerplanungslehre.¹⁸ Ziel der Steuerplanungslehre ist die Optimierung der (betrieblichen) Steuerbelastung durch das Aufstellen sowie die Anwendung von betriebswirtschaftlichen Entscheidungsmodellen bei der Steuerplanung.¹⁹ Solche Modelle basieren vorrangig auf rationalen Verhaltensannahmen und setzen monetäre

¹⁴ Vgl. Hechtner (2010), S. 1.

¹⁵ Vgl. exemplarisch Hundsdorfer/Sichtmann (2008).

¹⁶ Bzgl. der Übertragung von Methoden bspw. aus dem Bereich der Psychologie und des Marketings sei hier exemplarisch auf Hundsdorfer/Sichtmann (2008) und Blaufus et al. (2009) verwiesen.

¹⁷ So stammt bspw. die Conjoint-Analyse ursprünglich aus der Psychologie und findet schon seit längerem Anwendung im Bereich des Marketings. Vgl. Luce/Tukey (1964), Green/Srinivasan (1978).

¹⁸ Vgl. Hundsdorfer/Kiesewetter/Sureth (2008), S. 63f.

¹⁹ Vgl. König (2004), S. 260, König/Wosnitza (2004), S. 1.

Größen (bspw. den Kapitalwert oder den Endwert) als Zielgröße fest.²⁰ Die bei der Steuerplanung auftretenden Effekte lassen sich in Zeit-, Bemessungsgrundlagen-, und Steuertarifeffekte unterteilen.²¹ Unter den Begriff der Steuertarifeffekte lassen sich insbesondere auch Progressionseffekte der Besteuerung subsumieren, welche aus dem progressiv ausgestalteten deutschen Einkommensteuertarif resultieren. Bei dieser Form der Kategorisierung handelt es sich allerdings nicht um eine trennscharfe Abgrenzung.²² So ist insbesondere das Zusammenspiel verschiedener Effekte für eine umfassende Steuerplanung von Interesse. Bisher wurden Zeit- und Progressionseffekte sowie deren Zusammenspiel zahlreich in der betriebswirtschaftlichen Steuerlehre untersucht.²³ Hingegen fehlte für die mit der Unternehmensteuerreform 2008 eingeführte Thesaurierungsbegünstigung eine ausführliche Analyse dieses Zusammenspiels bei der Steuerplanung.²⁴ Dabei eignet sich diese Regelung besonders gut für eine Untersuchung dieses Zusammenspiels, da die Thesaurierungsbegünstigung dem Steuerpflichtigen erlaubt, über das Ausmaß seiner Progression in einem gewissen Rahmen selbst zu entscheiden. In einem ersten Schritt schließt dieser Beitrag somit diese Forschungslücke und unterzieht die Steuerplanung bei der Thesaurierungsbegünstigung einer umfassenden Analyse. Außerdem wird anhand dieses Beispiels gezeigt, wann die Vernachlässigung der Progression bei der Steuerplanung zu relevanten Fehlentscheidungen führt. Hierzu werden die Ergebnisse aus einem vereinfachten linearen Modell²⁵ mit den Ergebnissen eines komplexen Progressionsmodells verglichen. Die bei der Zielerreichung auftretenden Endvermögensdifferenzen geben anschließend Aufschluss über den isolierten Einfluss der Progression und lassen eine Schlussfolgerung über den Nutzen der Progressionsberücksichtigung in Steuerplanungsmodellen zu. Damit geht der Beitrag über die reine Analyse der Thesaurierungsbegünstigung hinaus und erweitert in einem zweiten Schritt den aktuellen Forschungsstand der Steuerplanungslehre.

²⁰ Vgl. *Wagner* (1995), S. 733.

²¹ Vgl. *Wagner* (1984), S. 211; *Wagner* (2005), S. 478-480. Auch *Hundsdoerfer* (2005), S. 635 sowie *Niemann/Kastner* (2009), S. 131 nutzen diese Aufteilung zur Zerlegung von steuerbedingten Gesamteffekten.

²² Vgl. *Hechtner* (2010), S. 23.

²³ Vgl. bspw. zu Fragen der Steuerbilanzpolitik: *Marettek* (1970), *Marettek* (1971); *Siegel* (1972); *Börner/Krawitz* (1977); *Siegel* (1980); *Marettek* (1980) sowie *Rückle* (1983).

²⁴ Insgesamt war diese Regelung seit ihrer Einführung bereits Gegenstand zahlreicher Untersuchungen. Vgl. exemplarisch *Homburg/Houben/Maiterth* (2008), *Houben/Maiterth* (2008a), *Houben/Maiterth* (2008b), *Kainz/Knirsch/Schanz* (2008). Eine Darstellung der optimalen Steuerplanung bei Inanspruchnahme der Thesaurierungsbegünstigung unter Berücksichtigung der Progression fehlte meines Erachtens allerdings bisher.

²⁵ Das Referenzmodell wurde aus *Homburg/Houben/Maiterth* (2008) entnommen.

Der letzte Beitrag, welcher sich mit fiskalpolitischen Effizienz der Spendenabzugsbegünstigung beschäftigt, lässt sich thematisch in das Forschungsfeld zwischen Betriebswirtschaftlicher Steuerlehre und Finanzwissenschaft einordnen. Eine klare Trennung zwischen dem Forschungsfeld der Betriebswirtschaftlichen Steuerlehre und der Finanzwissenschaft ist schwierig.²⁶ Während sich die Betriebswirtschaftliche Steuerlehre eher mit einzelwirtschaftlichen Problemen der Besteuerung beschäftigt, steht bei der Finanzwissenschaft mehr das Verhältnis zwischen dem Staat und den Bürgern im Mittelpunkt.²⁷ Dieses Abgrenzungskriterium ist allerdings keinesfalls trennscharf und die beiden Forschungsgebiete vermischen sich in jüngerer Zeit immer stärker.²⁸ Die in dem vorliegenden Beitrag untersuchte Fragestellung liegt ebenfalls auf dieser Schnittstelle. Es werden Spendenpreiselastizitäten für einzelne Steuerpflichtige abgeleitet, welche Aufschluss über die Bereitschaft dieser Steuerpflichtigen zur Leistung einer Spende geben. Damit werden einzelwirtschaftliche Entscheidungen von Individuen bzw. Haushalten untersucht. Dieses Vorgehen lässt sich somit noch in den Bereich der Betriebswirtschaftlichen Steuerlehre, speziell der empirischen Steuerwirkungslehre, einordnen.²⁹ Zur Beurteilung der fiskalpolitischen Effizienz dieser Begünstigungsnorm wird aber sowohl die Einnahmen- wie auch die Ausgabenseite des Staates betrachtet. Es wird untersucht, ob die Finanzierung besonders förderungswürdiger Zwecke eher privat durch Spenden oder besser direkt durch Staatsausgaben finanziert werden sollte. Damit fällt dieser Beitrag ebenfalls in das eher der Finanzwissenschaft zuzurechnende Forschungsfeld. Der Beitrag steht somit ebenfalls für die zunehmende Interdisziplinarität der Betriebswirtschaftlichen Steuerlehre. Der Schwerpunkt dieses Beitrags liegt meines Erachtens allerdings eher auf der Finanzwissenschaft.

Die Einzelbeiträge des zweiten Themenblocks decken somit ganz unterschiedliche Forschungsbereiche (Steuerplanungslehre und empirische Steuerwirkungslehre/Finanzwissenschaft) ab. Da bei beiden Beiträgen allerdings die Untersuchung einzelner steuerlicher Begünstigungsnormen im Mittelpunkt stehen, lassen sich diese Beiträge unter einen gemeinsamen Themenkomplex subsumieren. Im Folgenden werden die Einzelbeiträge

²⁶ Vgl. *Hechtner* (2010), S. 6.

²⁷ Vgl. *Fischer/Schneeloch/Sigloch* (1980), *Wöhe* (1988), sowie zur Finanzwissenschaft *Brümmerhoff* (2007).

²⁸ Weitere Abgrenzungskriterien finden sich mit weiteren Nachweisen bei *Hechtner* (2010), S. 8f.

²⁹ Ob Haushalte zum Untersuchungsgegenstand der Betriebswirtschaftlichen Steuerlehre gehören, ist in der Literatur allerdings umstritten. So grenzt bspw. *Schult* (2002), S. 4 die Untersuchung von Haushalten aus der Betriebswirtschaftlichen Steuerlehre aus. Eine andere Ansicht vertreten *Rose* (1970), S. 59 sowie *Hechtner* (2010), S. 14.

der beiden Themenkomplexe hinsichtlich ihrer Forschungsfrage, der angewendeten Methodik sowie der gewonnenen Erkenntnisse genauer dargestellt.

3. Forschungsfragen der Einzelbeiträge

3.1 Die gefühlte Steuerbelastung des Einkommens – Eine empirische Analyse für Erwerbstätige in Deutschland

Dieser Beitrag beschäftigt sich mit der Frage, wie das derzeitige deutsche Einkommensteuersystem von den Steuerpflichtigen wahrgenommen wird. Während in der bisherigen steuerwissenschaftlichen Literatur vorrangig davon ausgegangen wird, dass Steuerpflichtige ihre tatsächliche Steuerbelastung kennen und auch diese bei ihren ökonomischen Entscheidungen berücksichtigen, leitet dieser Beitrag eine subjektiv gefühlte Steuerbelastung ab. Das Wissen über die wahrgenommene Steuerbelastung ist von großem Interesse für die betriebswirtschaftliche Steuerforschung. So müssen Modelle und deren Annahmen, die der Vorhersage des tatsächlichen Verhaltens von Steuerpflichtigen in der Realität dienen, überdacht werden, wenn es eine Diskrepanz zwischen der tatsächlichen Steuerbelastung und der die Handlungen der Steuerpflichtigen beeinflussenden gefühlten Steuerbelastung gibt. Zusätzlich werden in diesem Beitrag sozio-demografische Determinanten bestimmt, die eine Fehleinschätzung der tatsächlichen Steuerbelastung beeinflussen. Zur Messung der gefühlten Steuerbelastung sowie ihrer Determinanten wird eine eigene Datenerhebung in Form einer Befragung durchgeführt.³⁰ Die Auswahl der Befragten richtet sich dabei nach einem Quotenplan des Statistischen Bundesamtes, um so eine repräsentative Zusammensetzung der Stichprobe hinsichtlich demografischer Merkmale entsprechend der Verteilung in der Grundgesamtheit der Erwerbstätigen in Deutschland zu erhalten.³¹ Neben der gefühlten Steuerbelastung wurde ebenfalls die von den Befragten jeweils als fair erachtete Steuerbelastung erhoben. Damit beantwortet dieser Beitrag zwei Forschungsfragen. Zum einen wird ermittelt, inwieweit Steuerpflichtige den Einkommensteuertarif überhaupt richtig wahrnehmen³² und zum anderen, ob dieser „gefühlte“ Einkommensteuertarif als gerecht empfunden wird.

Die Ergebnisse dieses Beitrags zeigen, dass die Befragten zu einem Großteil nicht in der Lage sind ihre Steuerbelastung richtig einzuschätzen, sondern dass die „gefühlte“ Steuerbelastung

³⁰ Ein ähnliches Vorgehen findet sich bereits bei *Schmölders* (1960), welcher ebenfalls die gefühlte Steuerbelastung mit Hilfe einer Befragung untersucht.

³¹ Zur Quotenstichprobe vgl. *Atteslander* (2008), S. 259f..

³² Ähnliche Untersuchungen finden sich auch für andere Länder, vgl. bspw. *Enrick* (1963, 1964), *Gensemer et al.* (1965), *Lewis* (1978), *Auld* (1979), *Fujii/Hawley* (1988), *Rupert/Fisher* (1995).

in den meisten Fällen von der tatsächlichen Steuerbelastung abweicht. Gerade bei hohen Einkommen wird die tatsächliche Steuerbelastung stark unterschätzt, wobei die Fehleinschätzung signifikant durch den Bildungsgrad des Befragten determiniert wird. Hinsichtlich der als gerecht empfundenen Steuerbelastung kann festgestellt werden, dass die tatsächliche Steuerbelastung für niedrige Einkommen eher als zu hoch und die Steuerbelastung für hohe Einkommen eher als zu niedrig empfunden wird. Auch diese Einschätzungen variieren unter den Befragten abhängig von Alter und Bildungsgrad.

Insgesamt weist dieser Beitrag auf eine große Diskrepanz zwischen der gefühlten, der tatsächlichen und der als fair erachteten Steuerbelastung hin. Dieses Ergebnis sollte innerhalb der Betriebswirtschaftlichen Steuerlehre dazu genutzt werden, Modellannahmen, die eine objektive Steuerwahrnehmung unterstellen und darauf aufbauend das Entscheidungsverhalten von Individuen prognostizieren wollen, zu überdenken.

3.2 The Influence of Tax Labeling and Tax Earmarking on the Willingness to Contribute – A Conjoint Analysis

Dieser Beitrag untersucht den Einfluss der Benennung (Labeling) von Steuern sowie deren Affektation auf die Bereitschaft der Steuerpflichtigen, diese Steuern trotz einer damit verbundenen Nettolohnminderung zu zahlen. Steuern werden häufig in kreativer Weise benannt. Beispiele dafür sind im Laufe der Steuergeschichte unter anderem der „Solidaritätszuschlag“, die „Bergarbeiterwohnungsbauabgabe“ oder die „Feuerschutzsteuer“.³³ Der Beitrag untersucht die Frage, inwieweit eine solche kreative Benennung ein geeignetes Instrument zur Senkung des Steuerwiderstandes darstellt bzw. die Akzeptanz und die Zahlungsbereitschaft bei den Steuerpflichtigen erhöht. Diese Fragestellung besitzt sowohl theoretische als auch steuerpolitische Relevanz.

In der Steuerwirkungslehre wird vorrangig unterstellt, Steuerpflichtige würden ihren Konsumnutzen und damit – bei gegebenem Arbeitsangebot – ihr Nettoeinkommen maximieren. Diese Annahme impliziert zugleich das Subziel der Steuerpflichtigen, ihre Steuerlast (allgemeiner: ihre Abgabenlast) zu minimieren, um so ein größtmögliches Konsumpotenzial zu erreichen. Insbesondere Untersuchungen zur Bereitschaft, Steuern zu hinterziehen, haben allerdings gezeigt, dass diese neoklassischen Annahmen in der Realität

³³ Vgl. *Schmölders* (1960), S. 83.

nicht immer erfüllt werden.³⁴ Viele Abgabepflichtige messen anscheinend Steuern und weiteren Abgaben sehr wohl einen positiven Nutzen bei, solange diese Abgaben ihrem persönlichen Gerechtigkeitsempfinden entsprechen und sie diese Abgaben als sinnvoll empfinden. Damit könnten Eigenschaften einer Abgabe, wie z. B. ihre Bezeichnung, die persönliche Einschätzung dieser Abgabe und das subjektive Belastungsempfinden beeinflussen. Wenn diese klassischen Grundannahmen bei der Analyse von steuerlichen Entscheidungswirkungen die Realität somit systematisch falsch einschätzen und Steuerpflichtige keineswegs stets bestrebt sind, ihr Nettoeinkommen zu maximieren, hat dies Auswirkungen sowohl auf die Entscheidungswirkungen einer Steuer als auch auf die wahrgenommene Lastenverteilung. Aus politischer Sicht könnten weiterhin mit dem Wissen über die Auswirkungen der Benennung sowie der Zweckbindung von Abgaben Ansatzpunkte identifiziert werden, um die Akzeptanz des Steuersystems in der Bevölkerung zu erhöhen.

Zu Beantwortung dieser Forschungsfrage wurden in diesem Beitrag Conjoint-Analysen mit 542 erwerbstätigen Personen in Deutschland durchgeführt. Dabei hatten die Teilnehmer die Möglichkeit, verschiedene Kombinationen von Abgaben in eine Rangfolge zu bringen.³⁵ Die hier zur Auswahl gestellten Abgaben waren der Beitrag zur Krankenversicherung, die zu zahlende Lohnsteuer sowie ein neuer Bildungszuschlag. Auf diesem Weg konnte ermittelt werden, wie die Benennung einer Abgabe (Labeling) sowie ihr Verwendungszweck (Affektation) auf die Akzeptanz dieser Abgabe in der Bevölkerung wirkt.

Das Ergebnis zeigt, dass sich in der Gesamtstichprobe gerade einmal ein Viertel der Probanden tatsächlich für das rationale Verhalten im Sinne der ökonomischen Theorie entschieden haben. Die Benennung der Abgabe (als „Lohnsteuer“, „Krankenversicherungsbeitrag“ oder „Bildungszuschlag“, Labeling-Effekt) und die Angabe über eine Zweckbindung der Abgabe „Bildungszuschlag“ (Affektationseffekt) beeinflussten signifikant die Bereitschaft zur Zahlung dieser Abgabe.

Für die Steuerwirkungsforschung bedeuten diese Resultate, dass Modelle, die die Anpassung von Entscheidungen an steuerrechtliche Rahmenbedingungen auf Basis der Rationalitätshypothese prognostizieren und weder die Benennung noch die Zweckbindung der Abgaben berücksichtigen, möglicherweise zu falschen Prognosen führen.

³⁴ Vgl. *Jonas/Heine/Frey* 1999 S. 164f. m.w.N.

³⁵ Zur Methode der Conjoint-Analyse vgl. *Luce/Tukey* (1964), *Green/Srinivasan* (1978) sowie *Backhaus et al.* (2009).

3.3 Steuerkomplexität und Arbeitsangebot - Eine experimentelle Untersuchung

Dieser Beitrag beschäftigt sich mit den Auswirkungen eines komplexen Steuertarifsystems auf die individuelle Arbeits-Freizeit-Entscheidung. Es wird ein betriebswirtschaftliches Experiment durchgeführt, um die in theoretischen Modellen häufig unterstellte Annahme zu überprüfen, dass Steuerpflichtige das jeweilige Steuersystem vollständig verstehen und somit in der Lage sind, steuerliche Regelungen und Vorschriften im Rahmen ihrer Steuerplanungs- und Steuergestaltungsmaßnahmen in rationaler Weise zu nutzen. Für dieses Vorgehen ist die Durchführung eines Experiments besonders geeignet, da dabei ein realitätsnahes Umfeld unter Setzung tatsächlicher monetärer Anreize geschaffen werden kann. In diesem Experiment wurden die Teilnehmer mit einer realen Arbeits-Freizeit-Entscheidung konfrontiert, bei welcher sie einen Bruttolohn für eine Arbeitstätigkeit erhielten, welche sie so lange ausführen konnten wie sie wollten. Alternativ konnte das Experiment jeder Zeit beendet werden (Freizeitalternative). Jeder Teilnehmer bekam zusätzlich ein für ihn geltendes Steuersystem, welches die auf seinen Bruttolohn anfallenden Steuerzahlungen in unterschiedlich komplexer Weise beschrieb. Die Komplexität des Steuersystems wurde dabei durch das Zusammenspiel verschiedener Steuersätze modelliert, welche additiv oder multiplikativ zu verknüpfen waren, was den Komplexitätsgrad jeweils beeinflusste. Diese Art der Modellierung lehnt sich an diverse Beispiele aus der Besteuerungspraxis an.³⁶

Die Forschungsfrage dieses Beitrags ist von theoretischem wie auch steuerpolitischem Interesse. Um aus theoretischen Modellen auf das tatsächliche Verhalten in der Realität schließen zu können ist es wichtig, die Verhaltensanpassungen der Steuerpflichtigen zu kennen, wenn diese mit einem komplexen Steuersystem konfrontiert werden. Die Annahme, dass Steuerpflichtige immer in der Lage sind, ihren Steuersatz richtig zu berechnen und auch dementsprechend zu handeln, muss überdacht werden, wenn die Zensiten ihre Steuerbelastung nur approximativ bzw. anhand von Vereinfachungsregeln oder Heuristiken ermitteln und ihr Verhalten somit gar nicht an dem tatsächlichen effektiven Gesamtsteuersatz ausrichten.³⁷

³⁶ So ist an dieser Stelle bspw. das Zusammenspiel zwischen Einkommensteuer und Solidaritätszuschlag bzw. der Kirchensteuer zu nennen. Im internationalen Kontext finden sich weiterhin Beispiele bei der Besteuerung internationaler Tätigkeit oder dem Zusammenspiel von „federal taxes“ und „local taxes“ in den USA.

³⁷ *Cogden et al.* (2009) führen bereits an, dass Steuerpflichtige eben nicht auf die tatsächliche Steuerbelastung, sondern auf die subjektiv wahrgenommene individuelle Belastung reagieren.

In politischen Debatten stehen Fragen der Steuervereinfachung ebenfalls immer wieder im Mittelpunkt.³⁸ Um über die Vor- und Nachteile einer Steuervereinfachung debattieren zu können, ist es allerdings wichtig zu wissen, welche Verhaltensanpassungen ein komplexes Steuersystem bei den Steuerpflichtigen hervorruft. Nur dann können das Ziel der Steuervereinfachung und dessen Folgen sinnvoll beurteilt werden.

Das zentrale Ergebnis dieses Beitrags ist, dass Steuerpflichtige auf komplexe Steuersysteme systematisch mit einer Einschränkung ihres Arbeitsangebots reagieren. Bei konstantem Netto- und Bruttolohn kann gezeigt werden, dass das Arbeitsangebot der Teilnehmer allein durch eine Steigerung des Komplexitätsgrades des Steuersystems signifikant sinkt. Weiterhin können Determinanten aufgezeigt werden, die diese Verhaltensanpassung beeinflussen. Die Hypothese, dass sich einzelne Steuerzahlungen in bzw. hinter der Komplexität des gesamten Steuersystems „verstecken“ lassen, kann im Ergebnis widerlegt werden.

3.4 Progressionseffekte und Varianten zur optimalen Steuerplanung bei der Thesaurierungsbegünstigung – Eine Abweichungsanalyse

Dieser Beitrag beschäftigt sich mit Progressionseffekten bei der Steuerplanung. Am Beispiel der Thesaurierungsbegünstigung des § 34a EStG wird untersucht, inwieweit die Vernachlässigung von Progressionseffekten zu Fehlentscheidungen bei der Steuerplanung führt. Des Weiteren wird der auftretende Fehler genauer quantifiziert und in einzelne Ursachen zerlegt. Die Thesaurierungsbegünstigung eignet sich besonders für diese Fragestellung, da diese dem Steuerpflichtigen erlaubt, in gewissen Grenzen über das Ausmaß seiner Progression selbst zu entscheiden. Außerdem wurden Progressionseffekte bei der Thesaurierungsbegünstigung bisher nur im Ansatz untersucht.³⁹ Als Referenzmaßstab dient das aus der Literatur bekannte lineare Optimierungsmodell von *Homburg/Houben/Maiterth*.⁴⁰ Die aus dem hier entwickelten Progressionsmodell abgeleiteten Ergebnisse werden mit diesem Referenzmodell verglichen, um den auftretenden Fehler durch die Vernachlässigung der Progression im Referenzmodell zu bestimmen.

³⁸ Vgl. dazu bspw. die Ziele der Koalitionsparteien im Bundestag, Koalitionsvertrag zwischen CDU, CSU und FDP, 17. Legislaturperiode, S. 12 ff.

³⁹ Vgl. bspw. Homburg/Houben/Maiterth (2008), Houben/Maiterth (2008a), Houben/Maiterth (2008b) sowie Kainz/Knirsch/Schanz (2008).

⁴⁰ Vgl. Homburg/Houben/Maiterth (2008).

Zur Lösung des Optimierungsproblems wurde ein Algorithmus entwickelt, der eine nicht lineare Optimierung im Rahmen eines zehnperiodigen Modells erlaubt.⁴¹ Als Zielgröße wurde dabei jeweils das Endvermögen mit und ohne Beachtung von Progressionseffekten maximiert, wobei die dabei auftretenden Endvermögensdifferenzen dann für verschiedene Einkommensklassen untersucht wurden. Im Ergebnis konnte gezeigt werden, dass es bei einer Vernachlässigung von Progressionseffekten zu Fehlentscheidungen bei der Steuerplanung kommen kann. Dabei tritt der Fehler insbesondere im Bereich mittlerer Einkommen auf, während bei niedrigen und sehr hohen Einkommen keine Fehler auftreten und eine detaillierte Steuerplanung somit unterbleiben kann. Weiterhin zeigt dieser Beitrag anschaulich, wie komplex die Steuerplanung unter Beachtung von Progressionseffekten werden kann. Mit diesem Ergebnis geht der vorliegende Beitrag über eine einfache Analyse der Thesaurierungsbegünstigung hinaus, da sich die gefundenen Ergebnisse leicht auf andere Probleme der Steuerplanung unter Berücksichtigung von Progressionseffekten übertragen lassen. Inwieweit sich eine detaillierte Steuerplanung unter Beachtung der entsprechenden Planungskosten lohnt, kann allerdings nur im Einzelfall entschieden werden.

3.5 Charitable Giving in the German Welfare State: Fiscal Incentives and Crowding Out

Spenden und Mitgliedsbeiträge zur Förderung steuerbegünstigter Zwecke können gemäß § 10b EStG in Grenzen als Sonderausgaben bei der Ermittlung der Steuerschuld abgezogen werden. Dieser Beitrag untersucht empirisch, ob die steuerliche Abzugsfähigkeit von Spenden aus fiskalpolitischer Sicht gerechtfertigt werden kann. Eine solche Rechtfertigung liegt dann vor, wenn der Verlust an Steuereinnahmen für den Staat durch die Gewährung des Spendenabzugs durch den Zuwachs an Spendenaufkommen selbst überkompensiert wird. In einem solchen Fall stellt sich der Staat besser, wenn er diese förderungswürdigen Zwecke nicht selbst finanziert, sondern einen Anreiz durch das Abzugssystem zur privaten Spendenfinanzierung setzt. Ist dies nicht gegeben, wäre es aus fiskalpolitischer Sicht allerdings vorzuziehen, den Spendenabzug abzuschaffen und die Finanzierung gemeinnütziger bzw. wohltätiger Zwecke direkt aus dem Staatshaushalt zu tätigen. Zusätzlich wird in diesem Beitrag untersucht, inwieweit es durch eine Erhöhung von staatlichen Ausgaben für wohltätige bzw. gemeinnützige Zwecke zu einem Spenden-Crowding-Out bei privaten Spendern kommt. Ein

⁴¹ Der Algorithmus wurde in Visual Basic for Applications umgesetzt und bedient sich der Solver-API der Lindo Suite, einer auf die Lösung nicht linearer Gleichungssysteme ausgerichteten Optimierungssoftware.

Spenden-Crowding-Out liegt dann vor, wenn private Spender einen Zuwachs an direkten staatlichen Ausgaben für förderungswürdige Zwecke zum Anlass nehmen, ihren privaten Spendenbeitrag zu reduzieren. Zur Untersuchung dieses Effekts wurde eine eigene Crowding-Out Variable aus den Daten des statistischen Bundesamtes generiert, welche für die einzelnen Bundesländern zahlreiche Ausgaben für gemeinnützige bzw. wohltätige Zwecke erfasst und somit einen guten Proxy zur Messung der entsprechenden Landesausgaben bietet.

Die Datengrundlage bildet die faktisch anonymisierte Lohn- und Einkommensteuerstatistik für die Veranlagungsjahre 1998, 2001 und 2004. Auf Basis dieser Daten werden die Spendenpreise der Steuerpflichtigen berechnet, welche die Grundlage für die Ermittlung der Spendenpreiselastizitäten darstellen. Weiterhin wird ein ökonomische Einkommen berechnet, welches das Konsumpotenzial der Steuerpflichtigen approximativ darstellt und als weitere Erklärungsvariable für das Spendenverhalten dient.

Zur Schätzung der Preis-, Einkommens- und Crowding-Out-Elastizitäten wird eine Quantilsregression⁴² durchgeführt, welche es erlaubt, Ergebnisse über das Spendenverhalten nicht nur allgemein, sondern gezielt für verschiedene Punkte der zugrunde liegenden Spendenverteilung zu gewinnen. Dazu werden die Spender entsprechend ihrer jeweiligen Spendenhöhe in verschiedene Quantile eingeteilt. Dieses Vorgehen wurde gewählt, da vermutet wurde, dass sich das Spendenverhalten in Abhängigkeit von der jeweiligen Spendenhöhe unterscheidet. Dieser Beitrag ist m.W. der erste, der diese Methode zur Untersuchung des Spendenverhaltens für Deutschland anwendet.⁴³

Ohne die Beachtung eines Spenden-Crowding-Outs deuten die gefundenen Ergebnisse darauf hin, dass die Mehrheit der Spender nicht preiselastisch auf eine Veränderung des Spendenpreises durch das Steuersystem reagiert und die steuerliche Spendenabzugsfähigkeit somit nicht gerechtfertigt werden kann. Erweitert man dieses Ergebnis allerdings um die gefundenen Crowding-Out-Effekte, so lässt sich die Spendenbegünstigung aus fiskalpolitischer Sicht rechtfertigen, da der Gesamteffekt ein preiselastisches Verhalten für alle Quantile der Spendenverteilung bestätigt. Neben potenziellen Crowding-Out-Effekten werden in diesem Beitrag zusätzlich weitere Gründe diskutiert, die ein Festhalten an der steuerlichen Spendenabzugsfähigkeit rechtfertigen können.

⁴² Detaillierte Informationen zur Quantilsregression finden sich bei *Chernozhukov/Hong* (2002).

⁴³ Ein ähnliches Vorgehen finden sich für Frankreich bei *Fack/Landais* (2010). Einen Überblick über die Ergebnisse ähnlicher Untersuchungen aus anderen Ländern und mit diversen anderen Methoden findet sich bei *Pelozo/Steel* (2005).

Literaturverzeichnis

- Atkinson, Anthony. B./ Stiglitz, Joseph E. (1976): The design of tax structure: direct versus indirect taxation, in: *Journal of Public Economics*, 6 (1), S. 55-75.
- Atteslander, Peter (2008): *Methoden der empirischen Sozialforschung*, 12. Auflage, Berlin 2008.
- Auld, Douglas A. L. (1979): Public Sector Awareness and Preferences in Ontario, in: *Canadian Tax Journal*, Jg. 27, H. 2, S. 172-183.
- Backhaus, Klaus/ Erichson, Bernd/ Plinke, Wulff/ Weiber, Rolf (2008): *Multivariate Analysemethoden*, 12. Auflage, Springer-Verlag, Berlin Heidelberg.
- Blaufus, Kay/ Bob, Jonathan/ Hundsdoerfer, Jochen/ Kieseewetter, Dirk/ Weimann/ Joachim (2009): It's All About Tax Rates - An Empirical Study of Tax Perception, arqus Diskussionsbeitrag zur Quantitativen Steuerlehre Nr. 106, November 2009.
- Börner, Dietrich/ Krawitz, Norbert (1977): *Steuerbilanzpolitik. Eine entscheidungsorientierte Analyse der Wahlrechte zur steuerlichen Gewinnermittlung*, Herne/Berlin.
- Brümmerhoff, Dieter (2007): *Finanzwissenschaft*, 9. Auflage, München 2007.
- Chernozhukov, Victor/ Hong, Han (2002): Three-step censored quantile regression and extramarital affairs, in: *Journal of American Statistical Association*, 97(459), 872-882.
- Congdon, William/ Kling, Jeffrey R./ Mullainathan, Sendhil (2009): Behavioral Economics and Tax Policy. *National Tax Journal*, 62, 375-386.
- Enrick, Norbert L. (1963): A Pilot Study of Income Tax Consciousness, in: *National Tax Journal*, Jg. 16, H. 2, S. 169-173.
- Enrick, Norbert L. (1964): A Further Study of Income Tax Consciousness, in: *National Tax Journal*, Jg. 17, H. 3, S. 319-321.

- Fack, Gabrielle/ Landaise, Camille (2010): Are tax incentives for charitable giving efficient? Evidence from France, in: *American Economic Journal: Economic Policy*, 2, S. 117–141.
- Fischer, Lutz/ Schneeloch, Dieter/ Sigloch, Jochen (1980): Betriebswirtschaftliche Steuerlehre und Steuerberatung – Gedanken zum 60jährigen „Jubiläum“ der Betriebswirtschaftlichen Steuerlehre, in: *Deutsches Steuerrecht 1980*, S. 699-705.
- Fujii, Edwin T./ Hawley, Clifford B. (1988): On the Accuracy of Tax Perceptions, in: *Review of Economics & Statistics*, Jg. 70, H. 2, S. 344-347.
- Gensemer, Bruce L./ Lean, Jane A./ Neenan, William B. (1965): Awareness of Marginal Income Tax Rates Among High-Income Taxpayers, in: *National Tax Journal*, Jg. 18, H. 3, S. 258-267.
- Green, Paul E./ Srinivasan, V. (1978): Conjoint Analysis in Consumer Research - Issues and Outlook. *Journal of Consumer Research* 5: 103-123
- Harberger, Arnold. C. (1964): The Measurement of Waste, in: *American Economic Review*, 54 (3), S. 58-76.
- Hechtner, Frank (2010): Eine theoretische und empirische Studie über Einkommensteuertarife aus Sicht der Wirtschaftswissenschaft – Progressionswirkungen der synthetischen Einkommensteuer, Schemulenbesteuerung und Vermeidung von Doppelbesteuerungen, Berlin.
- Homburg, Stefan/ Houben, Henriette/ Maiterth, Ralf (2008): Optimale Eigenfinanzierung der Personenunternehmen nach der Unternehmensteuerreform 2008/2009, in: *Zeitschrift für betriebswirtschaftliche Forschung*, 60. Jg., S. 29-47.
- Houben, Henriette/ Maiterth, Ralf (2008a): „Reichensteuer“ und Thesaurierungsbegünstigung versus 42%iger Spitzensteuersatz, in: *Finanz-Rundschau*, 63. Jg., S. 1044-1046.

- Houben, Henriette/ Maiterth, Ralf (2008b): Optimale Nutzung und Wirkungen von § 34a EStG, in: Steuer und Wirtschaft, 87. Jg., S. 228-237.
- Hundsdoerfer, Jochen (2005): Ist die Steuerbarwertminimierung auf die interpersonale Übertragung von Bemessungsgrundlagenteilen anwendbar?, in: Schneider, Dieter/ Rückle, Dieter/ Küpper, Hans-Ulrich/ Wagner, Franz (Hrsg.), Kritisches zu Rechnungslegung und Unternehmensbesteuerung, Festschrift zur Vollendung des 65. Lebensjahres von Theodor Siegel, Berlin 2005, S. 633-654.
- Hundsdoerfer, Jochen/ Kiewewetter, Dirk/ Sureth, Caren (2008): Forschungsergebnisse in der betriebswirtschaftlichen Steuerlehre – eine Bestandsaufnahme, Zeitschrift für Betriebswirtschaft, S. 61-139.
- Hundsdoerfer, Jochen/ Sichtmann, Christina (2007): Zur Gewichtung von Steuern in unternehmerischen Investitionskalkülen: Eine Conjoint-Analyse des Entscheidungsverhaltens selbständiger Ärzte, in: Betriebswirtschaftliche Forschung und Praxis, 604-621.
- Hundsdoerfer, Jochen/ Sichtmann, Christina (2009): The importance of taxes in entrepreneurial decisions: an analysis of practicing physicians' behavior, in: Review of Managerial Science 2009, S. 19-40.
- Jacobs, Otto (2004): Stand und Entwicklungstendenzen der Betriebswirtschaftlichen Steuerlehre, in: Steuer und Wirtschaft 2004, S. 251-259.
- Jonas, Eva/ Heine, Klaus/ Frey, Dieter (1999): Ein Modell der Steuerzufriedenheit – psychologische Grundlagen (un)ökonomischen Handelns, in: Fischer, Lorenz/ Kutsch, Thomas/ Stephan, Ekkehard (Hrsg.): Finanzpsychologie.
- Kainz, Robert/ Knirsch, Deborah/Schanz, Sebastian (2008): Schafft die deutsche oder österreichische Begünstigung für thesaurierte Gewinne höhere Investitionsanreize?, arqus Diskussionsbeitrag Nr. 41.
- Kirchler Erich (2007): The economic psychology of tax behavior. Cambridge University Press, New York.

- Koalitionsvertrag zwischen CDU, CSU und FDP (2009): Wachstum. Bildung. Zusammenhalt, 17. Legislaturperiode.
- König, Rolf (2004): Theoriegestützte betriebswirtschaftliche Steuerwirkungs- und Steuerplanungslehre, in: *Steuer und Wirtschaft*, S. 260-266.
- König, Rolf/ Wosnitza, Michael (2004): *Betriebswirtschaftliche Steuerplanungs- und Steuerwirkungslehre*, Heidelberg.
- Lewis, Alan (1978): Perceptions of Tax Rates, in: *British Tax Review*, Jg. 6, S. 358-366.
- Luce, R. Duncan/ Tukey, John W. (1964): Simultaneous conjoint measurement: A new type of fundamental measurement, in: *Journal of Mathematical Psychology* 1, S. 1-27.
- Marettke, Alexander (1970): Entscheidungsmodell der betrieblichen Steuerbilanzpolitik – unter Berücksichtigung ihrer Stellung im System der Unternehmenspolitik, in: *Betriebswirtschaftliche Forschung und Praxis*, 21. Jg., S. 7-31.
- Marettke, Alexander (1971): *Steuerbilanz- und Unternehmenspolitik*, Freiburg i. Br.
- Marettke, Alexander (1982): Die Stellung der Steuerplanung im Gesamtplansystem der Unternehmung, in: *Wirtschaftswissenschaftliches Studium*, 11. Jg., S. 19-25.
- Mirrlees, James A. (1971): An Exploration in the Theory of Optimum Income Taxation, in: *The Review of Economic Studies*, 38 (2), S. 175-208.
- Niemann, Rainer/ Kastner, Christoph (2009): Wie streitanfällig ist das österreichische Steuerrecht?, in: *Steuer und Wirtschaft*, 86. Jg., S. 128-138.
- Pelozo, John/ Steel, Piers (2005): The price elasticities of charitable contributions: a meta-analysis, in: *Journal of Public Policy and Marketing*, 24(2), S. 260–272.
- Ramsey, Frank P. (1927): A Contribution to the Theory of Taxation, in: *Economic Journal*, 37 (145), S. 47-61.

- Rose, Gerd (1970): Steuerberatung und Wissenschaft, in: Thoma, Gerhard/ Zacharias, O./ Niemann, Ursula (Hrsg.), Steuerberater-Jahrbuch 1969/70, Köln 1970, S. 31-70.
- Rückle, Dieter (1983): Normative Theorie der Steuerbilanzpolitik, Wien.
- Rupert, Timothy J./ Fischer, Carol M. (1995): An Empirical Investigation of Taxpayer Awareness of Marginal Tax Rates, in: Journal of the American Taxation Association, Jg. 17, H. 2, S. 36-59.
- Schmiel, Ute (2009): Forschungsziele der Betriebswirtschaftlichen Steuerlehre in der Kritik, in: Zeitschrift für Betriebswirtschaft, S. 1193-1214.
- Schmölders, Günter (1960): Das Irrationale in der öffentlichen Finanzwirtschaft, Rowohlt, Hamburg.
- Schneider, Dieter (1995): Betriebswirtschaftslehre, Band 1: Grundlagen, 2. Auflage, München 1995
- Schneider, Dieter (2002): Steuerlast und Steuerwirkung – Einführung in die steuerliche Betriebswirtschaftslehre, München 2002
- Schult, Eberhard (2002): Betriebswirtschaftliche Steuerlehre, 4. Auflage, München 2002.
- Siegel, Theodor (1972): Verfahren zur Minimierung der Einkommensteuer-Barwertsumme, in: Betriebswirtschaftliche Forschung und Praxis, 23. Jg., S. 65-80.
- Siegel, Theodor (1980): Auseinanderfallen von Steuerentstehung und Steuerzahlung bei der Steuerbilanzplanung, in: Zeitschrift für Betriebswirtschaft, 50. Jg., S. 377-386.
- Wagner, Franz (1974): Zum gegenwärtigen Forschungsprogramm der betriebswirtschaftlichen Steuerlehre, in: Der Betrieb 1974, S. 393-398.
- Wagner, Franz (2005): F.1. Besteuerung, in: Bitz, Michael/ Domsch, Michael/ Ewert, Ralf/ Wagner, Franz (Hrsg.), Vahlens Kompendium der Betriebswirtschaftslehre, Band 2, 5. Auflage, München 2005, S. 407-477.

- Wagner, Franz W. (1984): Grundfragen und Entwicklungstendenzen der betriebswirtschaftlichen Steuerplanung, in: Betriebswirtschaftliche Forschung und Praxis, 54. Jg., S. 201-222.
- Wagner, Franz W. (1995): Leitlinien steuerlicher Rechtskritik als Spiegel betriebswirtschaftlicher Theoriegeschichte, in: Elschen, Rainer/ Siegel, Theodor / Wagner, Franz (Hrsg.): Unternehmenstheorie und Besteuerung – Festschrift zum 60. Geburtstag von Dieter Schneider, Wiesbaden, S. 724-746.
- Wagner, Franz W. (2004): Gegenstand und Methoden betriebswirtschaftlicher Steuerforschung, in: Steuer und Wirtschaft 2004, S. 237-250.
- Wagner, Franz/ Schwenk, Anja (2003): Empirische Steuerwirkungen als Grundlage einer Reform der Gewinnbesteuerung – Ergebnisse aus den DAX 100-Unternehmen, in: Schwaiger, Manfred und Harhoff, Dietmar (Hrsg.): Empirie und Betriebswirtschaft, Entwicklungen und Perspektiven, Stuttgart 2003, S. 373-398.
- Westermann, Rainer (2000): Wissenschaftstheorie und Experimentalmethodik. Ein Lehrbuch zur Psychologischen Methodenlehre. Göttingen: Hogrefe.
- Wöhe, Günter (1988): Betriebswirtschaftliche Steuerlehre, Band I, 1. Halbband, München 1988.

Zusammenfassung

Bei der vorliegenden Arbeit handelt es sich um eine kumulative Dissertationsschrift, welche sich aus fünf Einzelbeiträgen zusammensetzt. Den gemeinsamen Forschungsschwerpunkt bildet die Untersuchung von Wahrnehmung und Wirkung der Ertragsbesteuerung sowie der (optimalen) Nutzung steuerlicher Begünstigungsnormen. Dabei ist die Arbeit wie folgt gegliedert: Zunächst wird im Rahmen eines Exposé eine Einordnung der Einzelbeiträge in den aktuellen Stand der betriebswirtschaftlichen Steuerforschung vorgenommen. Außerdem werden die Einzelbeiträge hinsichtlich Thematik, Methodik und Ergebnis kurz dargestellt. Im Anschluss daran folgen die fünf Einzelbeiträge.

Der erste Beitrag untersucht die Wahrnehmung der Besteuerung bei den Steuerpflichtigen. Es wird mit Hilfe einer Befragung untersucht, inwieweit die tatsächliche Steuerbelastung von der wahrgenommenen und der als gerecht erachteten Steuerbelastung bei den Zensiten abweicht. Die Ergebnisse zeigen, dass die Mehrheit der Befragten weder ihren Grenz- noch ihren Durchschnittssteuersatz richtig einschätzt. Insbesondere die Schulbildung der Befragten hat dabei einen signifikanten Einfluss auf die Abweichung zwischen gefühltem und tatsächlichem Steuersatz.

Der zweite Beitrag zeigt dann anhand einer Conjoint-Analyse, wie die Benennung und die Affektation von Abgaben dazu beitragen kann, die Zahlungsbereitschaft bei den Abgabepflichtigen zu erhöhen und somit den Steuerwiderstand zu senken. Es kann gezeigt werden, dass durch die Benennung und die Affektation einer Abgabe die Zahlungsbereitschaft signifikant beeinflusst werden kann.

Der dritte Beitrag untersucht im Rahmen eines Experiments unter Setzung tatsächlicher monetärer Anreize, welche Auswirkungen die Komplexität des Steuersystems auf die Arbeits-Freizeit-Entscheidung ausübt. Damit geht dieser Beitrag einen Schritt weiter und untersucht neben der Wahrnehmung komplexer Steuersysteme auch gezielt deren Entscheidungswirkungen. Das Ergebnis dieses Beitrags zeigt, dass das Arbeitsangebot der Teilnehmer signifikant abnimmt, wenn der Komplexitätsgrad des Steuersystems steigt.

Im vierten Beitrag wird dann eine modelltheoretische Analyse der Thesaurierungs-begünstigung (§ 34a EStG) unter besonderer Beachtung von Zins- und Progressionseffekten durchgeführt. Dieser Beitrag zeigt, wie komplex eine umfassende Steuerplanung in der Realität tatsächlich ist. Außerdem wird aufgezeigt, dass insbesondere für Bezieher mittlerer

Einkommen eine Vernachlässigung von Progressionswirkungen zu erheblichen Fehleinschätzungen bei der Steuerplanung führen kann.

Der letzte Beitrag greift dann als weitere steuerliche Begünstigungsnorm die Spendenabzugsfähigkeit (§ 10b EStG) heraus und untersucht empirisch, inwieweit sich diese Norm aus fiskalpolitischer Sicht rechtfertigen lässt. Das Ergebnis zeigt, dass unter Einbeziehung von Spenden-Crowding-Out-Effekten die steuerliche Abzugsfähigkeit aus fiskalpolitischer Sicht durchaus gerechtfertigt werden kann.

Summary

This cumulative doctoral dissertation consists of five essays. Common research aim of all essays is the analysis of tax perception and tax effects of income taxes as well as the analysis of the utilization of selected tax privileges. This thesis is organized as follows: First, the five theses will be classified regarding the current economic tax research. Furthermore, the topic, methodology and results of the individual essays will be presented. After this, follow the five individual essays.

The first article examines the perception of taxation on taxpayers. Using data from a survey, the paper studies perceived average and marginal personal income tax rates as well as the “fair” tax rates for different levels of income and compares them to actual tax rates. This study provides evidence that the perceived income tax rates significantly deviate from the objective tax rates for the majority of taxpayers. The degree of misperception can be partly explained by the individual's education level, income, and whether the individual included social security contribution in its income tax rate estimation. A comparison of the perceived tax rate with the tax rate that is perceived as fair shows that the perceived fairness depends on the individual's education level and age.

The second paper applies conjoint analysis to study the influence of tax labeling and tax earmarking on German taxpayers' willingness to contribute. It shows that labeling and earmarking effects can substantially increase taxpayers' willingness to contribute.

The third paper contains a controlled real-effort laboratory experiment to analyze the effects of complex tax systems on the individual work-leisure-decision. The results show that tax payers work supply decreases when the complexity of the tax system increases.

Using a mathematical optimization model, the fourth paper shows for the German tax privilege for retained earnings in which case and for which types of taxpayers the inclusion of progression effects in tax planning is appropriate. In result the inclusion of progression effects is especially important for medium income tax payers.

The fifth paper analyses tax privilege for charitable giving from a fiscal point of view. From a sample of German income tax returns, the elasticity of charitable giving relative to tax incentives, income, and government spending is estimated. Evaluating overall treasury efficiency, the tax deductibility of charitable giving fosters enough private giving to offset foregone tax revenues.

Lebenslauf

Der Lebenslauf ist in der Online-Version aus Gründen des Datenschutzes nicht enthalten.

Publikationsverzeichnis

Umsatzsteuerliche Behandlung ausländischer Gastwissenschaftler an deutschen Hochschulen, in: Umsatzsteuer-Rundschau 2009 (5), S. 145-148, (mit J. Hundsdorfer).

Belastungsverschiebungen durch das Familienleistungsgesetz - Kindergelderhöhung sowie Änderungen bei Kinderbetreuungskosten und haushaltsnahen Dienstleistungen, in: Finanz-Rundschau 2009 (2), S. 55-65, (mit F. Hechtner und J. Hundsdorfer).

Die Günstigerprüfung des Familienleistungsausgleichs nach dem zweiten Konjunkturpaket, in: Finanz-Rundschau 2009 (12), S. 573-578, (mit F. Hechtner).

Zur Bedeutung von Progressionseffekten für die Steuerplanung – Eine Analyse am Beispiel der Thesaurierungsbegünstigung, arqus-Diskussionspapier 101/2010, (mit F. Hechtner und J. Hundsdorfer).

The Name Game for Contributions – Influence of Labeling and Earmarking on the Perceived Tax Burden, FACTS Discussion Paper, School of Business & Economics, FU Berlin, Nr. 2010/12, (mit J. Hundsdorfer, K. Blaufus, D. Kiesewetter und J. Weimann).

Charitable Giving in the German Welfare State: Fiscal Incentives and Crowding Out, Economics Discussion Paper, School of Business & Economics, FU Berlin, Nr. 2010/30, (mit T. Bönke und N. Massarrat-Mashhadi).

Die gefühlte Steuerbelastung des Einkommens – Eine empirische Analyse für Erwerbstätige in Deutschland, in: Steuer und Wirtschaft 2010 (4), S. 337-345, (mit K. Blaufus, J. Bob, J. Hundsdorfer, D. Kiesewetter und J. Weimann).

Der steuerliche Lock-in-Effekt, in: Wirtschaftswissenschaftliches Studium 2011 (2), S. 96-99.

Progressionseffekte und Varianten zur optimalen Steuerplanung bei der Thesaurierungsbegünstigung – Eine Abweichungsanalyse, in: Schmalenbachs Zeitschrift für betriebswirtschaftliche Forschung 2011 (5), S. 214-239, (mit F. Hechtner und J. Hundsdorfer).

Besteuerung von Stückzinsen nach den Änderungen durch das JStG 2010, in: NWB Steuer- und Wirtschaftsrecht 2011 (7), S. 518-524, (mit F. Hechtner).

Charitable Giving in the German Welfare State: Fiscal Incentives and Crowding Out, in: Public Choice 2011 (online first), (mit T. Bönke und N. Massarrat-Mashhadi).

Steuerkomplexität und Arbeitsangebot – Eine experimentelle Analyse, FACTS Discussion Paper, School of Business & Economics, FU Berlin, Nr. 2011/13.

The Influence of Tax Labeling and Tax Earmarking on the Willingness to Contribute – A Conjoint Analysis, arqus-Diskussionspapier, August 2011, (mit J. Hundsdorfer, K. Blaufus, D. Kiesewetter und J. Weimann).

Anmerkung: Die grau hinterlegten Fachbeiträge sind Teil dieser Dissertation.

Überblick über die Einzelbeiträge

- Die gefühlte Steuerbelastung des Einkommens – Eine empirische Analyse für Erwerbstätige in Deutschland, in: *Steuer und Wirtschaft* 2010 (4), S. 337-345, (mit K. Blaufus, J. Bob, J. Hundsdorfer, D. Kiesewetter und J. Weimann).
- The Influence of Tax Labeling and Tax Earmarking on the Willingness to Contribute – A Conjoint Analysis, arqus-Diskussionspapier, August 2011, (mit J. Hundsdorfer, K. Blaufus, D. Kiesewetter und J. Weimann).
- Steuerkomplexität und Arbeitsangebot – Eine experimentelle Analyse, FACTS Discussion Paper, School of Business & Economics, FU Berlin, Nr. 2011/13.
- Progressionseffekte und Varianten zur optimalen Steuerplanung bei der Thesaurierungsbegünstigung – Eine Abweichungsanalyse, in: *Schmalenbachs Zeitschrift für betriebswirtschaftliche Forschung* 2011 (5), S. 214-239, (mit F. Hechtner und J. Hundsdorfer).
- Charitable Giving in the German Welfare State: Fiscal Incentives and Crowding Out, in: *Public Choice* 2011 (online first), (mit T. Bönke und N. Massarrat-Mashhadi).

Die gefühlte Steuerbelastung des Einkommens – Eine empirische Analyse für Erwerbstätige in Deutschland

Kay Blaufus, Jonathan Bob, Jochen Hundsdoerfer, Christian Sielaff, Dirk Kieseewetter
Joachim Weimann

Aus Urheberrechtsgründen ist dieser Beitrag in der Online-Version nicht enthalten. Der Aufsatz ist in *Steuer und Wirtschaft* 2010 (4), S. 337-345 erschienen, <http://www.otto-schmidt.de/>.

The Influence of Tax Labeling and Tax Earmarking on the Willingness to Contribute – A Conjoint Analysis*

Jochen Hundsdoerfer, Christian Sielaff, Kay Blaufus,
Dirk Kieseewetter, Joachim Weimann

August 11

Abstract

We apply conjoint analysis to study the influence of tax labeling and tax earmarking on German taxpayers' willingness to contribute. From a survey based sample we show that labeling and earmarking effects can substantially increase participants' willingness to contribute, which results in a considerable deviation from a pure consumption maximizing behavior. Furthermore, we give an explanation for this effect regarding socio-demographic attributes of German taxpayers. These results explain the variety in tax labels and provide implications for tax policy regarding further reforms of the tax and contribution system: "Labeling" and "earmarking" of contributions are important instruments in "selling" policies and increasing tax revenue.

Keywords

Behavioral Taxation · Tax Labeling · Tax Earmarking · Willingness to Contribute · Conjoint Analysis · Perceived Tax Burden

JEL Classification

D03 · H20 · H51 · H52 · K34

* This is a revised version of our earlier working paper: „*The Name Game for Contributions – Influence of Labeling and Earmarking on the Perceived Tax Burden*“, FACTS Discussion Paper, School of Business & Economics, FU Berlin, Nr. 2010/12.

Author of correspondence: Christian Sielaff, Department of Finance, Accounting and Taxation, Freie Universität Berlin, Garystraße 21, 14195 Berlin, Germany. E-mail: christian.sielaff@fu-berlin.de.

Contents

- List of Tables..... 3
- 1. Introduction 4
- 2. Literature Review 5
- 3. Conceptual framework 7
 - 3.1 Research hypothesis 7
 - 3.2 Method 10
 - 3.3 Sample 11
 - 3.4 Operationalization of the Hypotheses 14
- 4. Results 16
 - 4.1 Analysis of Ranking Behavior 16
 - 4.2 Influence of Labeling 18
 - 4.2.1 Education Allowance 18
 - 4.2.2 Health Insurance Premium 22
 - 4.3 Influence of Earmarking 23
 - 4.4 Influence of Individual Attributes 25
- 5. Conclusion..... 27
- References 29

List of Tables

Table 1 - Distribution of Quota Attributes	13
Table 2 - Contribution Combinations of Stimuli in Complete Design	15
Table 3 - Total Contributions and Net Income	16
Table 4 - Reference Sequence	17
Table 5 - Average Rank Placement of the Individual Stimuli	17
Table 6 - Part Worth Utilities and Relative Significance (Non-Earmarking Group).....	20
Table 7 - Sequences “Pro Education Allowance” with Non-Earmarking.....	21
Table 8 - Part Worth Utilities separated by Groups “Education Allowance” and “No Education Allowance”	22
Table 9 - Part Worth Utilities and Relative Significance (Earmarking Group).....	24
Table 10 - Sequence “Pro Education Allowance” with Earmarking	25
Table 11 - Explanation of the Labeling Effect with respect to individual attributes	26

1. Introduction

Throughout the world taxes are labeled in creative ways. For example, taxes were named as “solidarity surcharge” for the former Eastern part of Germany, as “miner residential construction charges” or as a “fire brigade tax”. Other examples outside the tax system are state lotteries that are promoted to “help the children” or welfare payments for promoting social fairness. This paper provides empirical evidence as to whether these instruments decrease taxpayers’ resistance to taxation and increase the acceptance of different taxes and contributions.

Studies on the willingness to evade taxes have shown that many taxpayers seem to attach positive connotations to taxes and further contributions as long as these contributions are consistent with their own sense of fairness and they find the contributions reasonable. Therefore, the attributes of a contribution such as its labeling and its possible earmarking may influence the individual evaluation of this duty and the subjective willingness to contribute. This study applies conjoint analysis to measure the influence of labeling and earmarking on the willingness to contribute. Conjoint analysis as an empirical method gives study participants the opportunity to rank different combinations of attributes. This method was used to determine how the naming of a contribution (labeling) as well as its designated use (earmarking) affects the acceptance of this contribution in the population.

This question is both of theoretical and political relevance. For theoretical research, it is important to analyze whether the traditional assumption of net income maximizing behavior holds for tax labeling and earmarking and which instruments can influence this behavior. From a political perspective, knowledge of the effects of labeling and earmarking contributions could be used to explain the variety of tax systems as well as to increase the acceptance of tax systems and, where applicable, to minimize tax avoidance strategies and evasion behavior.

This article is structured as follows: Section 2 presents a short literature review. Section 3 discusses the conceptual framework as well as the used method and the data. Section 4 presents the main results of the study. The final section concludes by summarizing and discussing key findings.

2. Literature Review

The starting point of the tax earmarking discussion in modern public finance is Buchanan (1963). He analyzes the economic effects of earmarked taxes and their influence on the individual's behavior within a theoretical model. He also discloses differences between earmarking and general-fund financing under restrictive model assumptions. Based on his work, Goetz (1968) develops a modified model and shows that the normative conclusions become less determinate under less restrictive assumptions. McCleary (1991) extends the theoretical models with World Bank case studies. He comes to the result that earmarking does not work very well in real settings. Durand, Klemmack and Lee Roff (1982) use cluster analyses to segment the respondents of a survey regarding their priorities for federal expenditures. They argue that the preferences for various expenditure programs differ on the basis of demographic and social psychological differences between the respondents and suggest the development of communication programs to inform the public about programs and goals in a more efficient manner. Based on these results, our study analyzes how the acceptance of a contribution is influenced by the information on its label and its earmarking. Furthermore, we analyze relevant demographical attributes to explain a potential increase in acceptance.

The effects of labeling and earmarking taxes are often discussed in papers that address the incentives of tax evasion. The focus of many studies is the impact of the public use of taxpayers' money on the willingness to evade taxes. Schmolders (1959), with further references, presents the results of a survey in Germany regarding general tax-mindedness and the moral assessment of tax dodging. He comes to the result that most respondents do not see tax dodging as a criminal activity and do not differ between taxes, voluntary contributions and fees. Kolm (1973) provides a basis for this in that he explicitly integrates the utility of public goods, which were financed through taxes involved in tax evasion decisions in an evasion model. Vogel (1974) determines through surveys that, in Sweden, the willingness to evade taxes decreased when the tax payment was accompanied by a certain consideration. Lewis (1979) reaches similar conclusions for Great Britain. Van de Braak (1983) offers a theoretical framework for the definition of tax resistance and discusses findings about tax compliance and tax mentality. Baldry (1986) shows that a difference between gambling and tax evasion is the existence of 'moral costs' in a situation deciding on the latter.

For a detailed survey about tax morale, see Torgler (2002) and Alm et al. (1995). Furthermore, experiments conducted by Spicer and Becker (1980) as well as Fortin et al.

(2007) show that the willingness to evade taxes decreases the higher the degree of the perceived tax justness and tax fairness. A model theoretical explanation for this is provided by Cowell (1992). Feld and Tyran (2002) conduct a laboratory experiment and come to the result that the higher the participants estimate the tax compliance of others to be, the higher their own tax morale. Alm, McClelland and Schulze (1992) come to the conclusion that individuals exhibit a great deal of diversity in their behavior towards taxation. They further determine that the main reasons for tax compliance are the overweighting of the low probability of audit and the taxpayers' valuation of public goods. Kim, Evans and Moser (2005) give a comprehensive explanation of the tax-reporting decisions and why they cannot be predicted accurately by conventional economic analysis. Harbaugh, Mayr and Burghart (2007) show that both required contributions and voluntary payments to a non-profit organization lead to similar brain activity amongst test persons as payments in their favor. In this observation, the authors see confirmation both for the motive of pure altruism and the motive of a "warm glow of giving", which is the benefit that the test person receives through his/her own voluntary donation.

A related body of literature attempts to measure subjective tax effects of contributions and subjective judgments of their fairness. According to Frey and Eichenberger (1996), the opportunity for taxpayers to participate in political tax questions is sufficient to increase tax satisfaction. These results are also verified by the results of Feld and Tyran (2002), who confirm a positive correlation between political participation and tax compliance. A detailed survey about tax fairness and tax justice is provided by Kirchler (2007) with further references.

Recently, the labeling and earmarking effect has become more and more important in both the economic and the psychological literature. Löfgren and Nordblom (2009) show that labeling can significantly affect attitudes towards a tax. According to their survey conducted in Sweden, a gasoline tax meets a stronger reluctance than the same tax when it is labeled as CO₂ tax on gasoline. Hardisty et al. (2010) analyze the effect of attribute framing on choice, labeling charges for environmental costs as either an earmarked tax or an offset. They find that the effect of attribute labeling depends on different political affiliations.

The general behavioral economics literature dismisses the usual assumptions made in economics about individual preferences. The basic deviations are that individuals act imperfectly rationally and that they hold non-standard preferences. Congdon et al. (2009) analyze the implications of behavioral economics for tax policy. They find that the welfare

consequences are ambiguous if taxpayers do not decide rationally. For example, people are not only self-interested but also care about the welfare of others (Andreoni and Miller 2002).

McCaffery and Baron (2006) with further references report findings of several experiments about perceptions of various tax-law design aspects. They come to the result that politicians can manipulate public opinion by tax system design, for example by using a “bonus” instead of a “penalty” (see Schelling 1984, p. 19 f.).

Based on the assumptions of behavioral economics literature this study explains the deviations from net income maximizing and pure self-interested behavior of taxpayers by analyzing the influence of tax labeling and tax earmarking. Our study shows that taxpayers’ behavior deviates systematically and considerably from consumption maximization.

This study builds upon previous results and enhances the status quo of research in two ways. On the one hand, we explicitly distinguish between a labeling and an earmarking effect on the willingness to contribute. On the other hand, we use conjoint analyses as a new method for this field of research to determine the willingness of German taxpayers to contribute. Conjoint analysis, well known in psychological and marketing research, is superior to survey based studies in the analysis of the willingness to contribute.

3. Conceptual framework

3.1 Research hypothesis

The reference point of our analysis is a taxpayer who aims to maximize his net consumption potential (called “reference” in the following). This net income maximizing behavior builds the initial point for evaluating the influences of labeling and earmarking on the willingness to contribute. The study does not aim to test the rational and self-interested “*homo oeconomicus*” model but rather to explore preferences for different tax policies and the influence of labeling and earmarking on that.

We consider three kinds of contributions: Income tax, health insurance premium and education allowance. The individual consumption potential is reduced by each of these contributions. In our analysis, each contribution may be low or high. Each possible combination of these contributions presents a treatment for the conjoint analysis. The specific contributions in our analysis have no direct monetary influence on the taxpayer’s net income, either because such direct considerations do not occur (income tax, education allowance) or

because their amount is explicitly assumed to remain unchanged (health insurance coverage). Thus, a pure net income maximizer would prefer low contributions.

According to this hypothesis, the name of a contribution (labeling) cannot influence the taxpayer's willingness to contribute because he aims to maximize his net income. Any potential statement of the designated use (earmarking) cannot affect the willingness to contribute as long as it deals with a matter in which the particular contribution amount does not directly increase the taxpayer's consumption potential.

The initial hypothesis under the assumption of a net income maximizer is stated as follows:

H_1 : For a given gross pay amount and given consideration from the system of contributions, taxpayers will choose the net income maximizing combination from a set of different combinations.

Now the contra thesis to H_1 must be developed. The subjective willingness to contribute, and thus the decision between combinations of contributions, could be influenced by the *labeling* of a contribution. Whereas in net income maximization, every type of contribution burden that causes a decrease in the individual consumption potential should be avoided, taxpayers could in reality factor the labeling of a contribution into their net burden calculations and with that, also into their decisions, even if their own consumption potential decreases in the process.

In addition to the egoistic behavior, the target function is now extended to include a limited egoistic or even an altruistic motive. A contribution labeled for example as a general education allowance could therefore receive wider acceptance among taxpayers than a contribution labeled as a tax if, for instance, taxes are associated with wastefulness by many taxpayers. The first contra thesis to H_1 reads as follows:

H_2 : The labeling of a contribution influences the willingness to contribute and therefore, the selection of taxpayers between combinations of contributions.

However, at this juncture, one must consider that the pure labeling of a contribution (e.g., education allowance) does not initially reveal anything about the concrete use of this contribution. The general labeling education allowance does not state whether the revenues from this contribution will be used, for example, to fund schools, universities, or general education projects such as community colleges, libraries, or museums. Certainly, this assumption holds only for a "one-shot-game". In the long run, taxpayers are able to observe if the money has been used according to the initial pledge of the government and this will

definitely influence the taxpayer's choice for the following rounds. Based on this study, further research should analyze the long run influence of labeling. In this study, only a "one-shot-game" is assumed.

The other tested contra thesis pertains to the appropriation (earmarking) of a contribution. Taxes are – in contrast to social contributions – usually not provided with appropriations (Non-Earmarking Principle, see e.g., McCleary 1991, p. 82). However, a non-enforceable political "pledge" could be given for a particular allocation of funds. In this way, appropriations would be possible for social contributions as well as for taxes.

It is questionable whether an appropriation can increase the acceptance of a contribution (McCleary 1991, p. 85). If taxpayers consider a particular cause for which a contribution is designated especially worthwhile, then they might be willing to accept a decrease in their consumption potential in order to support this cause. A similar behavioral pattern arises, for instance, concerning the willingness to give donations to non-profit organizations or particularly worthy causes. Accordingly, wage earners could, instead of choosing the alternative which offers them the highest net income, intentionally relinquish a portion of their consumption potential in order to support an allocation of funds that they consider worthwhile, without receiving a concrete service in return.

In this study, the study participants were divided into two groups to model the appropriation and the concrete use of the education allowance. One of these groups did not receive any detailed information regarding the use of the education allowance. The other group was informed about the concrete use of this contribution: Fifty percent of the education allowance was to be dispersed to schools and vocational schools, 25% to colleges and universities, and 25% to preschool education, since it was expected that the eligibility for the financing of this cause would, in general, be viewed positively by the population.

Two different conjoint analyses were conducted:

- In the first conjoint analysis, study participants were explicitly told that the tax revenue from the tested contribution (education allowance) was not earmarked. In this setting, solely a labeling effect was to be active.
- In the second conjoint analysis an earmarked education allowance was tested. In this case, both a labeling effect and definite earmarking were present.

This approach leads to the proposition of an additional sub-hypothesis as well as a contra hypothesis to H_1 :

H_3 : The earmarking of a contribution influences the willingness to contribute and therefore, the selection of taxpayers between combinations of contributions.

The two influential factors studied, labeling and earmarking, are not completely free from overlap, since the name of a contribution could be understood by taxpayers as a signal for the appropriation. However, it is possible to differentiate between a contribution with “lower” appropriation (solely by labeling of the contribution) and a contribution with “higher” appropriation (by a promised use): How does a “vague signal” for a possible appropriation (e.g., solidarity surcharge in Germany) act compared to clearly signaled earmarking, which is certain upon introduction of the contribution?

3.2 Method

The testing of the hypotheses mentioned above was carried out based on two conjoint analyses. A conjoint analysis consists of a procedure that was originally developed in the field of psychology (Luce and Tukey 1964) and later also found widespread use in market research (Green and Srinivasan 1978, pp. 103f.). In a conjoint analysis, study participants are assigned the task of evaluating various *bundles of attributes* (“stimuli”) according to their preferences (CONsider JOINTly), either by explicit evaluation (*rating*) or by indicating an order of precedence based on advantageousness (*ranking*). A bundle of attributes contains independent *factors* whereby each factor is described by a specific value called *level*. It is a multivariate technique and based on the premise that subjects evaluate the value of an object by combining the separate amounts of value provided by different attributes. The most fundamental concept and the conceptual basis for measuring value in conjoint analysis is the utility, a subjective judgment of preference unique to each individual (Hair et al. 2008, p. 266). The measured total utility of the various bundles of attributes are used to calculate part worth utilities and the relative importance of the single attributes. A conjoint analysis also consists of a decompositional procedure, which generally assumes an additive relationship of the part worth utilities with regard to the total utility (Hair et al. 2008). The advantage of the procedure is that study participants can holistically evaluate the objects to be analyzed, such that the demands on the study participants are lower and biases due to cognitively directed over/under weighting can be avoided. As a result, the conjoint analysis can be used to split up the total utility of a bundle of attributes into the part worth utilities of the individual attributes

and their relative weights. Since, in a conjoint analysis, the study participants have to give a simultaneous evaluation of different attributes, they must consider the same trade-off effects between the individual attributes that also exist in reality.¹

It would also be conceivable to directly question the study participants regarding willingness to contribute. However, the majority of study participants were not in a position to give consistent estimates of the burden which also take such trade-off effects into account. The study participants would probably tend to rate all attributes as especially important and therefore, overestimate the individual attributes in a bundle of attributes. The procedure of conjoint analysis was therefore chosen since it sets much lower demands on the study participants.

We consider that conjoint analyses are an applicable method for tax issues because of the advantages named above. In the field of marketing research, the conjoint analysis has been a standard method for several years. To this point, for tax issues, the application area of conjoint analysis is limited, but a conjoint analysis has already been used sporadically to measure tax effects. Milliron and Toy (1988) adopt conjoint analysis to research influential factors with regard to tax compliance. Hundsdoerfer and Sichtmann (2009) investigate the influence of taxation on investment decisions with the help of conjoint analysis. Blaufus and Ortlieb (2009) study the influence of tax complexity on employee decisions regarding company pension plans. To our knowledge, conjoint analysis has not yet been used to measure the influence of labeling and earmarking on the subjective burden of this contribution.

3.3 Sample

The data collection took place in December, 2008, and January, 2009, in the form of standardized face-to-face interviews with specially trained interviewers. In addition to the conjoint analysis, further questions were asked regarding demographic attributes, the general assessment of fiscal policy, and individual knowledge of German income tax law. The survey lasted an average of 20 minutes per study participant. Study participants were selected from among gainfully employed individuals, since these persons could be considered as being burdened by taxes and other contributions and thus, have experience with the German system of contributions. Gainfully employed individuals include, for purposes of this survey, all

¹ A *partition dependence effect* is avoided by using a complete stimuli design containing all possible combinations of attributes. See Fox et al. (2005, p. 539) for further details.

employees, public officials, and self-employed workers. Short interruptions to employment, for instance due to parental leave, maternity leave, etc. were not problematic.

The selection of study participants followed a quota schedule² which corresponded pro rata with the total working population in terms of the attributes of gender, age, education, and income level. The corresponding frequency in the population for the year 2006 was taken from the 2008 statistical almanac from the Federal Statistical Office in Germany, in which 37 million people comprised the working population in Germany. This was done to avoid bias from a non-representative composition of the sample since a representative random sample was not possible for financial reasons.

The conjoint analysis was conducted with a total of 542 study participants. This group was divided into two subgroups: 274 study participants (non-earmarking group) received no additional information about the appropriation of the contribution “education allowance”, whereas 268 study participants (earmarking group) received additional information regarding the concrete use of the contribution. Compliance with the quota schedule was statistically tested. No significant difference between sample and population could be detected at the 5% level of confidence, and the sample can therefore be seen as representative in regard to the studied traits. The majority of the interviews took place in Berlin and its surroundings. We controlled for differences between study participants from the former Western and Eastern Germany. More than twenty years after the end of the German division, we did not find differences between these groups in the data set. In addition, no significant difference could be identified between the two groups in the sample (non-earmarking group and earmarking group) with respect to the quota schedule at the 5% level of confidence.

The sample is composed of the individual attribute traits as follows:

² Quota samples do not strictly fulfill the requirements of a pure random selection. Nevertheless, they are the most widely used procedure in marketing research and consistently yield good results in comparative studies with pure random selection (see Green et al. 1988, pp. 325-327).

Table 1 - Distribution of Quota Attributes

Attribute	Value	Frequency	Percent (sample)	Percent (population)
Gender	Male	291	53.7 %	54.8 %
	Female	251	46.3 %	45.2 %
Age	Under 20 Years	20	3.7 %	3.5 %
	20 – 29 Years	93	17.2 %	17.5 %
	30 – 39 Years	127	23.4 %	24.0 %
	40 – 49 Years	152	28 %	29.6 %
	50 – 59 Years	115	21.2 %	20.5 %
	Over 60 Years	35	6.5 %	5.1 %
	No School Leaving Certificate	11	2.0 %	2.2 %
Education ³	Lower Secondary School Leaving Certificate	152	28.0 %	31.4 %
	Secondary School Leaving Certificate	148	27.3 %	26.8 %
	University-Entrance Diploma	86	15.9 %	14.5 %
	University Degree	101	18.6 %	16.0 %
Monthly Net Income	Other	44	8.1 %	9.1 %
	Under 1,000 Euro	164	30.3 %	32.7 %
	1,000 – 2,000 Euro	253	46.7 %	44.8 %
	2,000 – 3,000 Euro	82	15.1 %	14.6 %
	Over 3,000 Euro	38	7.0 %	8.0 %
	Not Stated	5	0.9 %	---

³ Lower Secondary School is the equivalent to the German “Hauptschule” while Secondary School is equivalent to the German “Realschule” and University-Entrance Diploma is equivalent to the German “Abitur”.

3.4 Operationalization of the Hypotheses

The research objects consist of a combination of contributions which must be paid by the study participants. The study participants were placed in a fictitious situation in which they were asked to rank⁴ combinations of contributions according to their preferences in determining the fiscal policy for an (unnamed) political party.⁵ The stimuli were presented to the study participants simultaneously on a table in front of them to make sure that the respondents fully understood and considered all the alternatives.

Every possible combination assumed a constant gross pay of 3,000 Euro. This gross pay was then reduced by three different contributions, i.e., deductions (factors). These deductions included a general income tax, a health insurance premium, and a newly introduced educational allowance. The factor levels were 700 or 740 Euro for income tax, 200 or 280 Euro for compulsory health insurance premiums, and 0 or 60 Euro for the education allowance. We explicitly advised the study participants that the health insurance premiums are independent of their actual received health insurance payments. Due to this advice our reference taxpayer will not pay a higher health insurance premium, independent of whether the taxpayer places some value on health benefits.

The net pay was also given for each stimulus to prevent arithmetic errors from distorting the study participants' results. There were a total of eight different possible combinations in the study (complete design), which were to be ranked by study participants during the course of the analysis according to their preferences. Each stimulus received a different level combination of the three factors (income tax, compulsory health insurance, education allowance).

The individual contribution combinations of the stimuli in the complete design are presented in Table 2.

⁴ This ranking is the most common valuation procedure after the rating scale. An overview of various procedures can be found in Green and Srinivasan (1978, pp. 104f).

⁵ To avoid a political party framing effect the name of the political party was unspecified.

Table 2 - Contribution Combinations of Stimuli in Complete Design

Stimulus	Income Tax	Education Allowance	Premium for Compulsory Health Insurance
A	low = 700 €	no	low = 200 €
B	low = 700 €	yes (60 €)	high = 280 €
C	high = 740 €	yes (60 €)	low = 200 €
D	high = 740 €	no	high = 280 €
E	high = 740 €	no	low = 200 €
F	low = 700 €	yes (60 €)	low = 200 €
G	high = 740 €	yes (60 €)	high = 280 €
H	low = 700 €	no	high = 280 €

The study participants were asked to assign the lowest rank in the sequence to the stimulus with their highest preference and the highest rank to the stimulus with their lowest preference. After a study participant had completed the sequence, the interviewer once again mentioned the possibility of making changes. When the study participant did not wish to make further changes, the final sequence was documented by the interviewer. The use of the education allowance was not stipulated in the first conjoint analysis (non-earmarking). In this case, explicit information was given to the study participants that tax revenue from the education allowance went into the general national budget and was not earmarked. In the second conjoint analysis (earmarking), study participants received detailed information regarding the use of the education allowance: Twenty-five percent of revenues from the education allowance were to be used for preschool education in nursery schools and day-care centers for children, 25% for improvements to colleges and universities, and 50% for improvements to schools and vocational schools. The sequences we collected can be used to test the initial hypothesis (hypothesis H₁). Each sequence created by a study participant can be compared with the reference sequence that a net income maximizing individual would choose in order to maximize his/her net income.

4. Results

4.1 Analysis of Ranking Behavior

The reference model of a net income maximizing individual provides a reference point for the examination of the empirically observed rank order of the preferred contribution system. This net income maximizing decision maker will always aim to minimize his/her contribution burden and therefore, maximize his/her net income.

Table 3 shows the total contributions and the maximum attainable net income of the individual stimuli.

Table 3 - Total Contributions and Net Income

Stimulus	Total Contributions	Net Income
A	900 €	2,100 €
B	1,040 €	1,960 €
C	1,000 €	2,000 €
D	1,020 €	1,980 €
E	940 €	2,060 €
F	960 €	2,040 €
G	1,080 €	1,920 €
H	980 €	2,020 €

If one ranks the stimuli according to the maximum attainable net income, this results in the following reference sequence

A>E>F>H>C>D>B>G,

since in this case, the stimulus with the highest net pay is ranked in first place and the ranking of the stimuli follows the total contributions.

In order to check how many study participants followed the example of the net income maximizing decision maker, Table 4 provides an overview of the number of net income maximizing sequences in the total sample.

Table 4 - Reference Sequence

Ranking	Number of Study Participants	Percentage of Study Participants
Reference	142	26.2 %
Other	400	73.8 %

Table 4 shows that the vast majority of study participants decided against a net income maximizing sequence. This result cannot be caused by arithmetic errors, since the exact net pay (gross pay – contributions) was indicated for every stimulus. Thus, net income maximization simply required the sorting of the stimuli according to net pay. Only 26.20% (142 individuals) of the study participants ranked all of the stimuli according to the reference rank order. A total of 73.80 % (400 individuals) instead chose another sequence. A t-test confirms that the deviations from the reference behavior are statistically significant (p-value < 0.01).

Table 5 shows the rank placement of the individual stimuli in the total sample:

Table 5 - Average Rank Placement of the Individual Stimuli

Stimulus	N	Median	Mean	Standard Deviation
Stimulus A (L-L-L)	542	1	2.63	2.090
Stimulus B (L-H-H)	542	7	5.38	2.067
Stimulus C (H-H-L)	542	5	4.01	1.645
Stimulus D (H-L-H)	542	6	6.29	1.426
Stimulus E (H-L-L)	542	3	3.79	2.049
Stimulus F (L-H-L)	542	3	2.75	1.641
Stimulus G (H-H-H)	542	8	6.56	1.966
Stimulus H (L-L-H)	542	4	4.58	1.557

The mean value is of particular interest. The mean indicates the average rank placement of a stimulus over the total sample. According to this average rank placement, we have the following sequence of stimuli:

A>F>E>C>H>B>D>G.

From Table 5, it becomes apparent that:

- Stimulus A has the lowest mean and therefore, on average, the best rank placement. This reflects the behavior of a net income maximizer. The same is true of stimulus G, which was valued as being the worst, both in net income maximization and on average in the sample.
- The sequence of stimuli according to the average of the study participants (mean) deviates from the sequence of stimuli in net income maximization. This leads to the assumption that the study participants do not follow the principle of net income maximization in their decision making.

Therefore, the behavior of the decision maker cannot be explained or predicted using the assumption of a *pure net income maximizer*. Thus, Hypothesis H₁ can be rejected. A further step must be used to test to what extent this deviation can be explained through the labeling of the contribution “education allowance” and the contribution “health insurance premium” (hypothesis H₂) as well as the earmarking of the contribution “education allowance” (hypothesis H₃).

4.2 Influence of Labeling

4.2.1 Education Allowance

According to hypothesis H₂, the labeling of a contribution leads to a change in the willingness to contribute and therefore, to a shift in contribution preferences. Thus, a contribution labeled as an “education allowance” should, according to hypothesis H₂, be evaluated differently by study participants from other contributions in a system of contribution combinations. Labeling the contribution as an “education allowance” affects the willingness to contribute twice. First the contribution is labeled as an “allowance” and not as a “tax”. Second, the label “education” suggests the designation for education purposes. In hypothesis H₂, we consider that this label increases the willingness to contribute. In order to eliminate bias in the effect of labeling, which is caused by the explicitly named appropriation of the education allowance (earmarking), the influence of labeling was first analyzed separately within the non-earmarking group.

Conjoint analysis makes it possible for each study participant to measure the part worth utilities of the individual attributes for his/her own sequence. Therefore the differences between the average of all ranks and the empirical ranks for each factor value were calculated.

This difference is the part worth utility of an individual factor level, which was standardized for comparison. The part worth utility is therefore a measure of the value that a study participant attaches to a factor level within the total attribute bundle. The estimation procedure is analogous to multiple regressions with dummy variables, but the calculations are done for each respondent separately, and it differs from other techniques dealing with relationships across all respondents. (For detailed information about the estimation technique see Hair et al. (2008 p. 269)). In this conjoint analysis, the part worth utility represents the utility that a study participant assigns to a contribution in a certain amount. The part worth utilities are used to cause a deviation in the individual willingness to pay a contribution. A high part worth utility denotes a high willingness to pay a contribution compared with the total willingness to pay the bundle of contributions.

As a means of comparison, the part worth utilities of the attributes could also be measured for a net income maximizer who ranks the stimuli solely according to the attainable net pay (Hundsdoerfer and Sichtmann 2009). In this way, we can test whether study participants assign a systematically higher value to the factor level “education allowance is levied” as it would be the case in the reference model. The correlation between the total utility of a stimulus and the individual part worth utility is – as is customary in conjoint analysis – assumed to be additive. This means that the sum of the individual part worth utilities of the attributes of a stimulus corresponds to the total utility.⁶ Table 6 shows the standardized⁷ part worth utilities and relative significance that result from the sample⁸ as well as the respective values in a reference sequence:

⁶ This conclusion is based on the assumption that the explanatory variables do not interact, see Hair et al. (2008).

⁷ The part worth utilities were standardized on a uniform scale of measurement and with a uniform zero-point.

⁸ The sample here refers only to the non-earmarking group.

Table 6 - Part Worth Utilities and Relative Significance (Non-Earmarking Group)

	Standardized Part Worth Utility (in %)		Relative Significance (in %)	
	Reference	Sample	Reference	Sample
Income Tax high	0.00	1.12	20.00	20.86
Income Tax low	20.00	19.74		
Education Allowance	0.00	12.92	33.33	40.21
No Education Allowance	33.33	27.29		
Health Insurance high	0.00	1.53	46.67	38.93
Health Insurance low	46.67	37.40		

In particular Table 6 shows that the part worth utilities for the education allowance in the sample deviate from the part worth utilities of the net income maximizing reference model. These results can be interpreted as follows: Whereas the characteristic “no education allowance” of a stimulus for our reference study participant accounts for 1/3 (33.33%) of the maximum total utility, this is only 27.29% for the average study participant.⁹ This result suggests that a contribution declared to be an education allowance finds wider acceptance than income tax.

A total of 76 of the 274 study participants in this part sample (non-earmarking group) preferred stimuli that levied the education allowance, so the part worth utility (mean) for the attribute “education allowance is levied” is at 12.92% and so clearly greater than for the reference point (part worth utility for this factor level = 0). Across all study participants, the difference between the part worth utility mean of the attributes “no education allowance” and “education allowance is levied” (27.29%-12.92% = 14.37%) showed that an increase in the education allowance from 0 to 60 € would burden the study participants to a lesser degree than would an increase in income taxes of 40 €(difference = 19.74% - 1.12% = 18.62%). This result shows that the willingness to pay the high education allowance is higher than the willingness to pay the high income tax. A t-test confirms the statistically significant deviation of the part worth utility for the factor level “education allowance is levied” from the reference value (p-value < 0.01).

⁹ For a rational, self interested study participant any label itself is without any importance, only the increase in net income for the characteristic “no education allowance” is important.

Thus, in the evaluation of the attribute “education allowance is levied”, many study participants apparently include not only the decrease in their consumption potential, but also attach positive importance to the promotion of education.

If study participants were willing to pay the education allowance and to relinquish a portion of their net income irrespective of how high the additional expenses were, then the sequence selected by these study participants would appear as follows:

F>C>B>G>A>E>H>D.

Here the sequence of the study participants corresponds to the reference sequence (net income maximizing) in regard to the attributes income tax and health insurance. However, the stimuli with the factor level “education allowance is levied” occupy the first four places. It’s a matter of an extreme case of endorsement for the education allowance. Table 7 reflects the frequency of this sequence compared to the reference sequence for the non-earmarking group (NE):

Table 7 - Sequences “Pro Education Allowance” with Non-Earmarking

Ranking	Number of Study Participants (NE)	Percentage of Study Participants (NE)	Total Sample
Reference	80	29.20 %	142 (26.20%)
Pro Education Allowance	14	5.11 %	47 (8.67%)
Other	180	65.69 %	353 (65.13%)
Sum	274	100.00 %	542 (100.00%)

Table 7 shows that a total of 14 out of 274 study participants are willing to pay the education allowance. A binomial test can also be used to show that the number of sequences carried out in this form present no statistically random result. Therefore, it can also be proven that a significant percentage of study participants intentionally chose the strategy “pro education allowance” instead of “contribution minimization” and therefore, chose the contributions with additional expenses. This clearly already appears sufficient to provide a contribution with a positive description. Hypothesis H₂ is therefore confirmed for the education allowance: The willingness to pay a contribution depends upon its labeling.

4.2.2 Health Insurance Premium

In addition to the previously established positive influence of the label “education allowance”, we analyzed whether the labeling of a contribution as “health insurance premium” influenced the willingness to contribute in contrast to a general income tax. In order to study this aspect, we divided the sample into two subgroups. We sub-divided each data set in such a manner as if the study participants had ranked each individual stimulus with an education allowance and without an education allowance. Therefore, two fictitious conjoint analyses per study participant were generated, each with four stimuli, in which only the attributes “income tax” and “health insurance premium” were varied and the education allowance remained constant (in one case 60 Euro and in the other case 0 Euro). So the labeling effect of the health insurance premium compared to the income tax could be isolated.

The calculated part worth utilities can be taken from Table 8.

Table 8 - Part Worth Utilities separated by Groups “Education Allowance” and “No Education Allowance”

Group Allocation	Income Tax low	Income Tax high	Health Insurance low	Health Insurance high
Education Allowance	31.86%	3.32%	61.75%	3.08%
No Education Allowance	37.64%	1.60%	56.15%	4.61%
For Comparison: Net Income Maximization	33.33%	0.00%	66.67%	0.00%

The willingness to pay the health insurance premium is noticeably higher in both groups than it would be under net income maximization.¹⁰ Thus, the term “health insurance premium” seems to be positive in comparison to income tax, or the study participants have – in contrast

¹⁰ The margin between low and high income tax is 40 Euro, which is half of the margin between the low and high health insurance premium. Accordingly, under net income maximization the part-worth utilities of the low income tax is half as high as for the low health insurance. For a comparison between the willingness to pay the health insurance premium and the income tax, the part worth utilities of the low income tax can be doubled. Then one recognizes – especially in the case without education allowance – the willingness to pay a higher health insurance premium is higher than to pay a higher income tax.

to the information that they received from us – associated a higher reward from compulsory health insurance with a higher health insurance premium.

When the study participants were confronted with an education allowance (group “education allowance”), then the willingness to pay a high health insurance premium is less than in the case in which no education allowance is collected. For study participants who are more strongly burdened by a higher health insurance premium than through a low one (the normal case), we see a difference of 4.35 percentage points.¹¹ The economic intuition behind this result is obvious: If study participants are willing to pay for an education allowance, then their inclination to “voluntarily” pay a higher health insurance premium decreases. This result was controlled with a Wilcoxon-Mann-Whitney Test and is significant on a 10% level of confidence. Therefore, a labeling effect can also be determined for the health insurance premium. This effect diminishes when an education allowance exists simultaneously.

4.3 Influence of Earmarking

It was shown above that the labeling of a contribution can change its acceptance among study participants. It is doubtful whether this effect is affected when study participants are explicitly informed of the purpose for which this contribution is to be used. The additional revenue generated by this increase in contributions should be explicitly assigned a previously determined use (earmarking). As mentioned above, a subgroup of study participants were given additional information for this purpose. Study participants in this subgroup were explicitly informed of the use of the education allowance.

The following part worth utilities resulted for the study participants who received this additional information. For comparison, the reference values as well as the results of the non-earmarking group are listed in Table 9.

¹¹ In order to be able to subtract the part worth utilities for the health insurance premium from one another, we excluded the study participants who had chosen a high health insurance premium over a low health insurance premium (“reversals”). For the low education allowance (N=504), one received a part worth utility of the low health insurance premium of 60.38%, and for the high education allowance (N=517), a part worth utility of 64.73%. The number of reversals for a low education allowance (542-504=38) is higher than for a high education allowance (542-517=25), and this also supports the results.

Table 9 - Part Worth Utilities and Relative Significance (Earmarking Group)

	Part Worth Utility			Relative Significance		
	Reference	Non-Earmark	Earmark	Reference	Non-Earmark	Earmark
Income Tax high	0.00 %	1.12 %	0.76 %	20.00 %	20.86 %	21.28 %
Income Tax low	20.00 %	19.74 %	20.52 %			
Education Allowance	0.00 %	12.92 %	24.54 %	33.33 %	40.21 %	41.82 %
No Education Allowance	33.33 %	27.29 %	17.28 %			
Health Insurance high	0.00 %	1.53 %	1.38 %	46.67 %	38.93 %	36.90 %
Health Insurance low	46.67 %	37.40 %	35.67 %			

In order to test whether there is a significant difference between the distribution within the earmarking and non-earmarking groups with regard to the acceptance of the education allowance, a Wilcoxon-/Mann-Whitney-U-Test was conducted. At a level of confidence less than 0.001%, the test showed that there is a significant shift in the distribution function between the two groups. In particular, the mean value of the part worth utility in the non-earmarking group is 12.9% and in the earmarking group 24.5%. As such, the appropriation leads to an average increase in the part worth utility of an education allowance of 11.6%. With that, the individual willingness to contribute for the education allowance is significantly higher if the study participants were informed about the specific appropriation for the contribution. Therefore hypothesis H₃ is confirmed: The appropriation of a contribution has an influence on the willingness to pay this contribution. The average willingness to pay the contribution “education allowance” increases further if the study participants are convinced of its validity not only through the labeling of the contribution, but also through its earmarking.

In our study, the effect is so strong that – with earmarking – the education allowance no longer presents a net burden. This is shown in Table 9: The mean value of the part worth utilities of the high education allowance is more than 7 percentage points *higher* than the mean value of the part worth utilities of the low education allowance. Therefore, the high education allowance creates a higher total utility on average in the “earmarking” group than the low education allowance.

Table 10 serves to clarify how many study participants in the earmarking group (E) consciously chose the high education allowance, but otherwise behaved as a net income maximizer.

Table 10 - Sequence “Pro Education Allowance” with Earmarking

Ranking	Number of Study Participants (E)	Percentage of Study Participants (E)	Total Sample
Reference	62	23.13 %	142 (26.20 %)
Pro Education Allowance	33	12.31 %	47 (8.67 %)
Other	173	64.55 %	353 (65.13 %)
Sum	268	100.00 %	542 (100.00%)

If one compares these results with the values in Table 7, it becomes clear that, in the earmarking group, a strikingly larger number of study participants chose the “pro education allowance” but otherwise chose the reference sequence (12.31%). This also confirms the hypothesis that the willingness to pay can be increased through the earmarking of a contribution. Thus, hypothesis H₃ is also confirmed: In addition to the labeling of a contribution, the willingness to pay a contribution can be increased further if the contribution is earmarked for a special use, which appears to be seen as especially worthwhile by numerous study participants.

4.4 Influence of Individual Attributes

In order to test to what extent demographic data can explain the results, different non-parametric test methods were carried out in which the dependent variable consisted of the part worth utility of a levied education allowance. Nonparametric test methods are appropriate because of the quality of the data. We separated the sample into different classes with respect to the education level, age, gender, children and income. We then tested each class against the other for differences with respect to their part worth utility for the high education allowance. As test methods the Wilcoxon-Mann-Whitney-Test for two class comparisons and the Kruskal-Wallis-Test for testing more than two classes are applied. Table 11 provides an overview of the studied influence factors.

Table 11 - Explanation of the Labeling Effect with respect to individual attributes

Test classes	Test method	Results
Earmarking vs. Non-Earmarking	Wilcoxon-Mann-Whitney-Test	p<0.001***; Earmark: Mean: 12.92%; 95% Konf.: 10.1-15.75%; Non-Earmark: Mean: 24.54%; 95% Conf: 21.07-28.00%
Education level (at least University-Entrance Diploma vs. Secondary School Leaving Certificate and Lower Secondary School Leaving Certificate) ¹²	Kruskal-Wallis-Test (and pairwise Wilcoxon-Mann-Whitney-Tests)	p=0.029**
Income level (<1000; -2000; -3000; >3000)	Kruskal-Wallis-Test	not significant
Age (<=20; -30; -40; -50; -60; >60)	Kruskal-Wallis-Test	not significant
Gender	Wilcoxon-Mann-Whitney-Test	not significant
Children	Wilcoxon-Mann-Whitney-Test	not significant

Other than for the attribute earmarking, no significant influence of the demographic attributes on the willingness to pay the education allowance with the exception of the education level can be determined. Solely the attribute “Lower Secondary School” allows the conclusion to be drawn at a 5% level of confidence that students at a lower secondary school tend to be less willing to voluntarily pay for education than students with at least a University-Entrance Diploma.¹³ It is conspicuous that the test for the class “children” is not significant: We expected that those who have children may regard an education allowance as a positive benefit, which affect the perceived consumption possibilities for the family. However, no significant differences between taxpayers with and without children regarding the acceptance of the “education allowance” were found. We conducted the same tests for the dependent variable high health insurance premium, but without any significant results.

¹² We excluded the class “No school leaving certificate” because of the small sample size and the class “Others” because of the lack of interpretability.

¹³ This was confirmed via pairwise Wilcoxon-Mann-Whitney-Tests. Only the test between the class “at least University-Entrance Diploma” and “Lower Secondary School Leaving Certificate” shows a significant result at a 5% level of confidence (p-value: 0.012).

5. Conclusion

The purpose of this article was to test whether labeling and earmarking are useful instruments in increasing the willingness to pay a contribution. The model of a net income maximizer served as a reference, under which individuals attempt to avoid all contributions that do not provide a concrete reward. In order to test this question empirically, we applied the instrument of conjoint analysis. We polled 542 gainfully employed individuals in Germany. The task of the study participants was to rank various combinations of contributions. From these rankings, we derived part worth utilities and the relative significance of the attributes of the individual contributions, which we could then compare with part worth utilities and the relative significance of a fictitious net income maximizer.

In the total sample, only one-fourth of the study participants actually chose the net income maximizing behavior described often by economic theory. The initial hypothesis that taxpayers always and exclusively strive to maximize their net pay and therefore, their consumption potential, must be rejected. The labeling of the contribution (as “income tax”, “health insurance premium”, or “education allowance”) and the description (earmarking) of the appropriation of the “education allowance” influence the willingness to pay this contribution. In order to distinguish the labeling effect and the earmarking effect from one another, study participants were divided into two subgroups, of which only one received additional information about the allocation of revenues from the education allowance. For the group that received no additional information regarding the allocation of funds, a statistically significant and economically relevant difference could be identified between the reference and the empirically proven sequencing behavior (labeling effect). It became clear that some study participants assigned a positive part worth utility to a high, but not concretely, earmarked education allowance although this is associated with a decrease in their net income. The labeling of a contribution therefore has an influence on its acceptance. A similar, yet weaker, effect in the same direction could be observed for the contribution “health insurance premium”. An appropriation of the contribution “education allowance” significantly strengthened this effect. On average, the study participants no longer described this contribution as a burden (negative net burden).

For tax effects research, these results indicate that models that predict the conformance of decisions on general fiscal conditions based on the net income maximizing hypothesis and which fail to take into account either the labeling or the appropriation of contributions, may

lead to inaccurate predictions. Whether our conclusions for working persons also apply to firms remains an open research question.

We suppose that the results found here are of great interest for future reforms to the tax and contribution systems. The labeling of a contribution and especially the definition of the use of the tax revenue generated led to an increase in the acceptance and willingness to pay among working individuals. An increase in acceptance likely also leads to diminished evasion and avoidance behavior and thus, to lower transaction costs and tax collection costs. Furthermore, our results approve that “labeling” and “earmarking” are important instruments for politicians in “selling” their tax or contribution policies.

There are numerous links between the current study and possibilities for further research. For instance, the question arises as to what extent the labeling and appropriation of individual types of contributions can also lead to negative perceptions and therefore, to a decrease in the willingness to contribute. When interpreting our results, it should be considered that this study dealt solely with a pure preference sequence, which had no economic consequences for the study participant. There is a need for further research which validates these results in connection with actual monetary incentives. As previously mentioned, the potential to transfer the results for contributions that are levied at a corporate level also remains to be tested. Another research question for further studies is to analyze our results in a multiple shot setting in order to test for taxpayers’ “learning effects”.

References

- Alm J, McClelland GH, Schulze W (1992) Why do people pay taxes?. *Journal of Public Economics* 48: 21-38.
- Alm J, Sanchez I, Juan A de (1995) Economic and noneconomic factors in tax compliance. *KYKLOS* 48: 3-18.
- Andreoni J, Miller JH (2002) Giving according to GARP: An experimental test of the consistency of preferences for altruism. *Econometrica* 70 (2): 737-753.
- Baldry JC (1986) Tax evasion is not a gamble. A report on two experiments. *Economic Letters* 22: 333-335.
- Blaufus K, Ortlieb R (2009) Is simple better? A conjoint analysis of the effects of tax complexity on employee preferences concerning company pension plans. *Schmalenbach Business Review* 61 (1): 60-83.
- Braak H van de (1983) Taxation and tax resistance. *Journal of Economic Psychology* 3: 95-111.
- Buchanan JM (1963) The economics of earmarked taxes. *The Journal of Political Economy* 71 (5): 467-469.
- Congdon WJ, Kling JR, Mullainathan S (2009) Behavioral economics and tax policy. *National Tax Journal* 62 (3): 375-386.
- Cowell FA (1992) Tax evasion and inequity. *Journal of Economic Psychology* 13: 521-543.
- Durand RM, Klemmack DL, Lee Roff L (1982) Citizen preferences for the allocation of federal tax dollars: A segmentation approach. *Journal of Marketing & Public Policy* 1: 169-180.
- Feld LP, Tyran, JR (2002) Tax evasion and voting: An experimental analysis. *KYKLOS*. 55: 197-222.
- Fortin B, Lacroix G, Villeval MC (2007) Tax evasion and social interactions. *Journal of Public Economics* 91: 2089-2112.
- Fox CR, Lieb DS, Ratner RK (2005) How subjective grouping of options influences choice and allocation: Diversification bias and the phenomenon of partition dependence. *Journal of Experimental Psychology* 134 (4): 538-551.

- Frey BS, Eichenberger R (1996) To harmonize or to compete? That's not the question. *Journal of Public Economics* 60: 335-349.
- Goetz CJ (1968) Earmarked taxes and majority rule budgetary processes. *The American Economic Review* 58 (1): 128-136.
- Green PE, Srinivasan V (1978) Conjoint analysis in consumer research - Issues and outlook. *Journal of Consumer Research* 5: 103-123.
- Green PE, Tull DS, Albaum G (1988) *Research for marketing decisions*. 5th edn. Prentice-Hall, Englewood Cliffs.
- Hair JF, Black WC, Barry JB, Anderson RE (2008) *Multivariate data analysis*, 7th edn. Pearson Education, New York.
- Harbaugh WT, Mayr U, Burghart DR (2007) Neural responses to taxation and voluntary giving reveal motives for charitable donations. *Science* 316: 1622–1625.
- Hardisty DJ, Johnson EJ, Weber EU (2010) A dirty word or a dirty world? Attribute framing, political affiliation, and query theory. *Psychological Science* 21 (1): 86-92.
- Hundsdoerfer J, Sichtmann C (2009) The importance of taxes in entrepreneurial decisions: An analysis of practicing physicians' behavior. *Review of Managerial Science* 3: 19-40.
- Kim CK, Evans JH, Moser DV (2005) Economic and equity effects on tax reporting decisions. *Accounting, Organizations and Society* 30: 609-625.
- Kirchler E (2007) *The economic psychology of tax behavior*. Cambridge University Press, New York.
- Kolm SC (1973) A note on optimum tax evasion. *Journal of Public Economics* 2: 265-270.
- Lewis A (1979) An empirical assessment of tax mentality. *Public Finance* 34: 245-257.
- Löfgren Å, Nordblom K (2009) Puzzling tax attitudes and labels. *Applied Economics Letters* 16: 1809-1812.
- Luce RD, Tukey JW (1964) Simultaneous conjoint measurement: A new type of fundamental measurement. *Journal of Mathematical Psychology* 1: 1-27.
- McCaffery EJ, Baron J (2006) Thinking about tax. *Psychology, Public Policy, and Law* 12 (1): 106-135.

- McCleary W (1991) The earmarking of government revenue - A review of some world bank experience. *World Bank Research Observer* 6 (1): 81-104.
- Milliron VC, Toy DR (1988) Tax compliance: An investigation of key features. *Journal of the American Taxation Association* 9: 84-104.
- Schelling TC (1984) *Choice and consequence*. Harvard University Press, Cambridge.
- Schmölders G (1959) Fiscal psychology: A new branch of public finance. *National Tax Journal* 12 (4): 340-345.
- Spicer MW, Becker LA (1980) Fiscal inequity and tax evasion: An experimental approach. *National Tax Journal* 33 (2): 171-175.
- Torgler B (2002) Speaking to theorists and searching for facts: Tax morale and tax compliance in experiments. *Journal of Economic Surveys* 16: 657-683.
- Vogel J (1974) Taxation and public opinion in Sweden: An interpretation of recent survey data. *National Tax Journal* 27: 499-513.

Steuerkomplexität und Arbeitsangebot - Eine experimentelle Analyse

Christian Sielaff*

August 2011

Abstract

Die vorliegende Arbeit untersucht den Einfluss von komplexen Steuertarifsystemen auf die Arbeits-Freizeit-Entscheidung von Individuen. Dazu wird ein wirtschaftswissenschaftliches Experiment mit 96 studentischen Teilnehmern durchgeführt. Es wird eine realitätsnahe Arbeits-Freizeit-Entscheidung unter Setzung tatsächlicher monetärer Anreize modelliert, um so den Einfluss verschieden komplexer Steuertarifsysteme auf das individuelle Arbeitsangebot der Experimentteilnehmer abzuleiten. Die Modellierung der verschiedenen Komplexitätsstufen erfolgt dabei durch das Zusammenspiel mehrerer Steuersätze, wie es auch in der Besteuerungspraxis zu finden ist. Es kann gezeigt werden, dass bei konstantem Netto- und Bruttolohn das Arbeitsangebot der Teilnehmer allein durch eine Steigerung des Komplexitätsgrades des Steuertarifsystems signifikant sinkt. Die Fragestellung ist sowohl für die theoretische Forschung als auch für zukünftige Steuerreformdebatten mit dem Ziel der Steuervereinfachung von Interesse.

Keywords

Steuerkomplexität · Arbeits-Freizeit-Entscheidung · Experimentelle Steuerforschung · Real-Effort-Experiment · Behavioral Taxation

JEL classification

C91 · H24 · H31

* Dipl.-Ök. Christian Sielaff, Lehrstuhl für Betriebswirtschaftliche Steuerlehre, Fachbereich Wirtschaftswissenschaft, Freie Universität Berlin, Email: christian.sielaff@fu-berlin.de

Inhaltsverzeichnis

1. Einleitung.....	4
2. Literaturüberblick.....	6
3. Spezifikation der Forschungsfrage und Experimentdesign	9
3.1 Spezifikation der Forschungsfrage.....	9
3.2 Experimentdesign.....	12
4. Ergebnisse und Diskussion	16
5. Zusammenfassung und Ausblick	33
Anhang	35
Literaturverzeichnis	39

Tabellenverzeichnis

Tabelle 1 - Steuertarifsysteme.....	14
Tabelle 2 - Gesamte Arbeitsleistung im Experiment.....	16
Tabelle 3 - Produktivitäten im Zeitablauf	17
Tabelle 4 - Arbeitsleistung nach Gruppen	19
Tabelle 5 - Paarweise Gruppenvergleiche	21
Tabelle 6 - Produktivitäten nach Zeitintervall und Steuertarifsystem	22
Tabelle 7 - Arbeitszeit und Arbeitsleistung in korrigierter Form	23
Tabelle 8 - Paarweise Gruppenvergleiche in korrigierter Form	24
Tabelle 9 - Fragen zum Experiment.....	26
Tabelle 10 - OLS Regression, Abhängige Variable: Arbeitszeit, Erweiterte Spezifikation	29
Tabelle 11 - OLS Regression, Abhängige Variable: Arbeitszeit, Basisspezifikation.....	30
Tabelle 12 - OLS Regression, Abhängige Variable: Arbeitsleistung, Erweiterte Spezifikation	31
Tabelle 13 - OLS Regression, Abhängige Variable: Arbeitsleistung, Basisspezifikation.....	32

Abbildungsverzeichnis

Abbildung 1 - Produktivitäten im Zeitablauf	18
Abbildung 2 - Arbeitszeit und Arbeitsleistung	20

1. Einleitung

Die Komplexität des deutschen Einkommensteuersystems steht immer wieder im Mittelpunkt zahlreicher politischer wie auch wissenschaftlicher Diskussionen. Thematisiert werden dabei insbesondere Fragen der Steuervereinfachung, welche regelmäßig von verschiedenen Seiten gefordert werden.¹ Zu welchen konkreten Verhaltensreaktionen auf Seiten der Steuerpflichtigen ein komplexes Steuertarifsystem führt, bleibt bei diesen Debatten allerdings häufig außer Acht. Der vorliegende Beitrag widmet sich deshalb dieser Frage. Es soll anhand einer experimentellen Untersuchung gezeigt werden, welchen Einfluss die Komplexität eines Steuertarifsystems auf Arbeitsangebotsentscheidungen eines Steuerpflichtigen ausübt.

Die Reaktion der Steuerpflichtigen auf Steuerkomplexität ist von zentraler Bedeutung, um über die Notwendigkeit und die Folgen einer Komplexitätsreduzierung innerhalb des deutschen Steuertarifsystems sinnvoll debattieren zu können, bzw. um über die Umsetzung von Steuervereinfachungen zu entscheiden. Die Durchführung eines Experiments erscheint dabei vor dem Hintergrund dieser Fragestellung als besonders geeignet, da durch ein Experiment ein realitätsnahes Umfeld geschaffen werden kann, welches erlaubt, das Verhalten und die Reaktionen von Steuerpflichtigen unter kontrollierten Bedingungen zu beobachten. Aus diesen experimentellen Ergebnissen können Prognosen über die Auswirkungen einer Steuervereinfachung abgeleitet werden. Insbesondere das Setzen tatsächlicher monetärer Anreize erlaubt eine angemessene Modellierung einer realen Arbeits-Freizeit-Entscheidung und eine Untersuchung des Einflusses einer tatsächlich zu zahlenden Steuerbelastung.

Die Untersuchung ist somit sowohl für die theoretische betriebswirtschaftliche Forschung, insbesondere für die steuerliche Verhaltensforschung (Behavioral Taxation), wie auch für die aktuelle politische Diskussion von großer Relevanz.

In der steuerökonomischen Modelltheorie wird im Allgemeinen unterstellt, dass Steuerpflichtige das jeweilige Steuertarifsystem vollständig verstehen und somit in der Lage sind, steuerliche Regelungen und Vorschriften im Rahmen ihrer Steuerplanungs- und Steuergestaltungsmaßnahmen

¹ Vgl. dazu bspw. die Ziele der Koalitionsparteien im Bundestag, Koalitionsvertrag zwischen CDU, CSU und FDP (2009) sowie den Entwurf eines Steuervereinfachungsgesetzes 2011, BT-Drucks. 17/5125. Ein aktueller Reformentwurf zur Erneuerung und Vereinfachung des Steuerrechts findet sich bei *Kirchhof* (2011). Dazu kritisch: *Wagner* (2011) sowie *Bach* (2011).

in rationaler Weise zu nutzen.² Aus diesen Modellen wird dann auf das tatsächliche Verhalten in der Realität sowie ggf. auf die damit verbundenen gesamtwirtschaftlichen Auswirkungen geschlossen. Sollte ein Großteil der Steuerpflichtigen aber aufgrund der Komplexität des Steuertarifsystems nicht in der Lage sein, steuerliche Sachverhalte in entsprechender Weise zu erkennen und so die damit verbundenen Planungs- und Gestaltungsmöglichkeiten zu nutzen, dann ist die Erklärungs- und Vorhersagekraft einer solchen Modelltheorie entsprechend niedriger. An dieser Stelle knüpft der vorliegende Beitrag an. Es soll empirisch untersucht werden, welche Verhaltensreaktionen Steuerpflichtige zeigen, wenn das zugrunde liegende Steuertarifsystem aufgrund der hohen Komplexität nicht verstanden wird. Diese Frage ist von großem theoretischen Interesse für die Steuerforschung, da die theoretischen Annahmen über die Wirkung der Besteuerung korrigiert werden müssen, wenn Steuerpflichtige nicht in der Lage sind, die Besteuerung korrekt in ihre ökonomischen Entscheidungen einfließen zu lassen, sondern ihr Verhalten vielmehr an der individuell wahrgenommenen Besteuerung und damit bspw. an Vereinfachungsregeln bzw. Heuristiken ausrichten.³

Weiterhin leistet diese Untersuchung einen wichtigen Beitrag zur steuerpolitischen Diskussion. Die Koalitionsparteien im Bundestag haben sich in ihrem Koalitionsvertrag darauf geeinigt, das Steuertarifsystem zu vereinfachen.⁴ Dieses Vorgehen steht im Gegensatz zur bisherigen Steuerpolitik, welche die Komplexität des Einkommensteuersystems in den letzten Jahren durch immer neue Sondertarifvorschriften wie die Einführung der Abgeltungsteuer oder die Thesaurierungsbegünstigung erhöht hat.⁵ Insgesamt existieren neben der Grundtarifvorschrift des § 32a EStG noch fünf weitere einkommensteuerliche Tarifnormen, die die Komplexität des Einkommensteuersystems wesentlich erhöhen.⁶ Um aus politischer Sicht zu entscheiden, welche Steuergesetzgebung vorzuziehen ist, sollten die politischen Entscheidungsträger auch die Auswirkungen steuerlicher Komplexität bzw. deren Vereinfachung bei den Zensiten berücksichtigen. Sollte neben der tatsächlichen Höhe der Besteuerung auch dem Komplexitätsgrad des Steuertarifsystems eine Bedeutung bei der Entscheidungsfindung seitens der Steuerpflichtigen zufallen, so muss dieser Effekt im

² So unterstellt bereits *Ramsey* (1927), dass Steuerpflichtige auf Steueränderungen in der gleichen Weise reagieren wie auf Preisänderungen. Diese Annahme wird auch in zahlreichen weiteren Beiträgen innerhalb der wirtschaftswissenschaftlichen Forschung übernommen, bspw. *Harberger* (1964), *Mirrlees* (1971), *Atkinson/Stiglitz* (1976).

³ Zur Anwendung von Heuristiken vgl. *Gigerenzer/Gaissmaier* (2011).

⁴ Vgl. Koalitionsvertrag zwischen CDU, CSU und FDP (2009).

⁵ Vgl. beispielsweise zur Abgeltungsteuer: *Hechtner/Hundsdoerfer* (2009), S. 41, oder zur Thesaurierungsbegünstigung: *Houben/Maiterth* (2008) sowie *Hechtner/Hundsdoerfer/Sielaff* (2011).

⁶ Vgl. *Hechtner* (2010), S. 27f.

Rahmen der Steuergesetzgebung berücksichtigt werden.⁷ Deshalb ist die hier untersuchte Fragestellung ebenfalls aus steuerpolitischer Sicht von Interesse.

Der vorliegende Beitrag gliedert sich wie folgt. Im zweiten Kapitel wird zunächst auf den aktuellen Stand der Literatur eingegangen. Kapitel 3 spezifiziert die untersuchte Forschungsfrage und beschreibt den Aufbau des Experiments. Kapitel 4 stellt die Ergebnisse des Experiments dar und diskutiert mögliche Ursachen und Implikationen. Kapitel 5 schließt diesen Beitrag mit einem zusammenfassenden Fazit sowie einem Ausblick auf zukünftigen Forschungsbedarf und geht auf Grenzen der Untersuchung ein.

2. Literaturüberblick

Die vorliegende Untersuchung knüpft insbesondere an die bisherige Literatur zu Wahrnehmung und Wirkung von komplexen Steuertarifsystemen an. *Cogden et al.* (2009) führen an, dass Steuerpflichtige nicht auf tatsächliche Steuersätze reagieren, sondern auf die individuell wahrgenommenen. Da gerade mit steigender Komplexität eines Steuertarifsystems der Spielraum für solche „Steuersatzinterpretationen“ steigt, scheint es naheliegend zu sein, diesem „interpretierten“ Steuersatz mehr Beachtung bei der Untersuchung von Steuerwirkungen zu schenken. Bisher wurde der Einfluss von Steuern auf ökonomische Entscheidungen in verschiedenen Studien untersucht, welche sich hinsichtlich der Modellierung des Steuertarifsystems, der genauen Fragestellung und der angewandten Methode stark unterscheiden.

So untersuchen *Gamage, Hayashi und Nakamura* (2010) im Rahmen eines Experiments den Einfluss der Präsentation des Nettoeinkommens auf das Arbeitsangebot. Bei diesem Experiment werden Probanden gebeten, sich nach Vorgabe eines bestimmten Nettolohns zu entscheiden, ob diese innerhalb eines bestimmten Zeitintervalls arbeiten wollen oder in dieser Zeit lieber einer vorgegebenen Freizeitbeschäftigung nachgehen wollen. Sie kommen zu dem Ergebnis, dass die Probanden eine höhere Arbeitsbereitschaft zeigen, wenn ihnen ein hoher Bruttolohn abzüglich einer

⁷ Der Komplexitätsgrad eines Steuertarifsystems könnte bspw. dazu führen, dass die Zensiten ihre Steuerbelastung anders wahrnehmen, als diese tatsächlich ist. Kleine Einzelsteuern könnten somit durch den Gesetzgeber im „Steuerwirrwarr“ versteckt werden und zu einer Differenz zwischen tatsächlicher und wahrgenommener Belastung führen, was wiederum den Steuerwiderstand beeinflussen kann. Zum subjektiven Belastungsgefühl vgl. *Schmölders* (1960), S. 80f. Vergleichbare Ansätze finden sich in der wirtschaftswissenschaftlichen Literatur bereits bei der Diskussion um die „Unmerklichkeit“ der Besteuerung. Hierbei werden die Vorteile einer indirekten Besteuerung im Vergleich zu einer direkten Besteuerung thematisiert, wobei ebenfalls die Senkung des Belastungsgefühl durch die Unmerklichkeit indirekter Steuern herausgestellt wird. Vgl. *Keller* (1968), S. 110; *Neumark* (1970), S. 37f.; *Tipke/Lang* (2010), § 8, Rn. 11. Dieses Vorgehen kann vom Grundsatz her auch auf Komplexitätswirkungen übertragen werden.

Steuer gezahlt wird anstatt eines niedrigen Nettolohns zzgl. einer Bonuszahlung. Hinsichtlich der Präsentation bzw. Darstellung von Steuern zeigen weiterhin *Chetty, Looney und Kroft* (2009) in einem Experiment, dass Steuern nur dann „richtig“ verarbeitet werden, wenn diese für den Steuerpflichtigen überhaupt erkennbar („salient“) sind. Auch das dient als Anknüpfungspunkt der vorliegenden Untersuchung, da die über mehrere Steuersätze modellierte Komplexität ebenfalls einen „salience effect“ darstellen kann. So werden die Teilnehmer ggf. durch eine der Komplexität geschuldeten längeren Beschreibung des Steuertarifsystems eher für die Steuerzahlung sensibilisiert.

Blumkin, Ruffle und Ganun (2008) modellieren im Rahmen eines Experiments ebenfalls eine Arbeits-Freizeit-Entscheidung, wobei die Probanden das verdiente Geld für reale Güter ausgeben müssen. Sie untersuchen dabei, welchen unterschiedlichen Einfluss eine Einkommensteuer und eine Konsumsteuer auf das Arbeitsangebot haben und kommen dabei zu dem Ergebnis, dass eine Einkommensteuer die Arbeitsbereitschaft stärker senkt als eine äquivalente Konsumsteuer.

Sillamaa (1999) untersucht im Rahmen einer Replikationsstudie zu *Swenson* (1988) den Einfluss verschiedener linearer Steuersätze auf das Arbeitsangebot und bestätigt das aus der Theorie bekannte Ergebnis, dass das Arbeitsangebot mit steigenden Steuersätzen und damit sinkendem Nettolohn sinkt.

Fochmann et al. (2010) beschäftigen sich mit den Auswirkungen eines Bruttolohneffekts bei der Arbeits-Freizeit-Entscheidung. Sie zeigen, dass das Arbeitsangebot von Probanden signifikant höher ist, wenn bei gleichem Nettolohn der Bruttolohn, aber dementsprechend auch der geltende Steuersatz, höher ist. Diesen Effekt führen *Fochmann et al.* (2008) auf eine „gross-wage illusion“ zurück.

Mit dem Einfluss der Komplexität auf die Wahrnehmung von Steuern beschäftigen sich *Chetty und Saez* (2009). Diese zeigen am Beispiel des Earned Income Tax Credit (EITC) in den USA, dass ein zusätzliches Informationsangebot zu Verhaltensanpassungen bei den Steuerpflichtigen führt, und dass der EITC zu kompliziert ist, um von den Steuerpflichtigen richtig wahrgenommen zu werden.

De Bartolome (1995) untersucht ebenfalls im Rahmen eines Experiments den Einfluss von Steuerkomplexität, indem er die Komplexität durch die unterschiedliche Präsentation des gleichen Steuertarifsystems variiert. *De Bartolome* (1995) kann zeigen, dass Steuern unabhängig von der Präsentationsform von den Probanden berücksichtigt werden, diese allerdings dazu neigen, den Durchschnittssteuersatz anstelle des Grenzsteuersatzes zu benutzen, wenn der Grenzsteuersatz nicht explizit bei der Darstellung des Steuertarifsystems ausgewiesen wird. Ein ähnliches Vorgehen findet sich bei *Rupert und Wright* (1998). Diese kommen zu dem Ergebnis, dass Steuerpflichtige

ihre Investitionsentscheidungen umso besser treffen, je transparenter ihnen der entsprechende Grenzsteuersatz präsentiert wird.

Ein weiteres Experiment zur Steuerkomplexität findet sich bei *Rupert et al.* (2003). Diese modellieren die Komplexität des Steuertarifsystems durch Abzüge von der Bemessungsgrundlage, welche dazu führen, dass die Berechnung des effektiven marginalen Steuersatzes komplexer wird. Bei diesem Experiment kann gezeigt werden, dass die Probanden seltener eine optimale Investitionsentscheidung treffen, wenn die Komplexität des Steuertarifsystems hoch ist. *Boylan* und *Frischmann* (2006) zeigen, ebenfalls anhand eines Experiments, dass ein komplexes Steuertarifsystem zu einer Beeinträchtigung der Markteffizienz führen kann.

Insbesondere die Wahrnehmung und Einschätzung von Grenzsteuersätzen wurde in der Literatur häufig behandelt. So finden sich dazu diverse Untersuchungen etwa bei *Gensemer et al.* (1965), *Lewis* (1978), *Fujii* und *Hawley* (1988), *Rupert* und *Fisher* (1995) sowie im Rahmen einer Conjoint-Analyse bei *Hundsdoerfer* und *Sichtmann* (2007). Diese Studien kommen größtenteils zu dem Schluss, dass Steuerpflichtige nicht in der Lage sind, ihren Grenzsteuersatz richtig zu berechnen und diesen somit für ökonomische Entscheidungen zu nutzen. Deshalb zeigen auch diese Untersuchungen, welche Folgen ein komplexes Steuertarifsystem und die damit verbundenen Fehleinschätzungen haben.

Ein anderes Vorgehen wählen *Blaufus* und *Ortlieb* (2009), welche zur Untersuchung von steuerlichen Komplexitätswirkungen auf die Methode der Conjoint-Analyse zurückgreifen. Steuerkomplexität wird hier anhand der Zeit, welche Steuerpflichtige brauchen, um ein Steuertarifsystem zu verstehen, gemessen. Die Steuertarifsysteme selbst unterscheiden sich dabei anhand der Länge der den Probanden vorgelegten Beschreibungen sowie durch die Anzahl der darin verwendeten Fachtermini. *Blaufus* und *Ortlieb* (2009) zeigen, dass mit steigender Komplexität immer weniger Steuerpflichtige die Steuerzahlungen bei ihren Entscheidungen berücksichtigen.

Das Verhältnis von Steuerkomplexität und Steuerehrlichkeit wurde im Rahmen eines Experiments von *Bizer* (2004, 2008) untersucht. *Bizer* zeigt, welchen Einfluss Unsicherheiten über die Höhe des zu versteuernden Einkommens auf die Bereitschaft zur Steuerhinterziehung ausüben. Er kommt zu dem Ergebnis, dass Komplexität Steuerehrlichkeit zerstört und Steuerpflichtige mehr Steuern hinterziehen, wenn die Höhe ihres zu versteuernden Einkommens unsicher ist. Ähnliche experimentelle Untersuchungen zum Verhältnis von komplexitätsbedingter Unsicherheit über die Steuerzahlung und Steuerehrlichkeit finden sich auch bei *Beck, Davis* und *Jung* (1991) sowie *Alm, Jackson* und *McKee* (1992).

Die hier durchgeführte Untersuchung baut auf den bisherigen experimentellen Studien auf und erweitert diese wie folgt. Zum einen wird die Arbeits-Freizeit-Entscheidung anders als in der bisherigen Literatur bekannt modelliert. Die oben aufgeführten Untersuchungen, mit Ausnahme von *Fochmann et al.* (2010), modellieren die Freizeitalternative bspw. durch das Hören von Musik, das Spielen von Computerspielen oder den Konsum von Freizeitgütern, was nur im Ansatz eine realistische Freizeitalternative darstellt. In der vorliegenden Untersuchung ist diese Entscheidung so ausgestaltet, dass die Probanden die entsprechende Arbeitstätigkeit jederzeit beenden und nach Hause gehen können.⁸ Dies entspricht eher der Entscheidungssituation in der Realität.

Wie oben bereits beschrieben, unterscheidet sich dieser Beitrag zum anderen durch die Modellierung der Steuersatzkomplexität und klammert dabei die bereits mehrfach untersuchten Effekte durch die Verwechslung von Grenz- und Durchschnittssteuersätzen aus. Der Beitrag widmet sich somit möglichen Komplexitätseffekten durch die Interaktion verschiedener linearer Steuertarife.

Außerdem werden tatsächliche Einkommenseffekte in diesem Experiment ausgeblendet. Im Gegensatz zu *Fochmann et al.* (2010) werden alle Probanden mit dem gleichen Bruttolohn konfrontiert, was eine „*gross-wage illusion*“ ausschließt. Auch der Einfluss des Nettolohns wird ausgeblendet, da sich die einzelnen Teilnehmergruppen nur hinsichtlich der Steuersatzkomplexität unterscheiden, der effektive Steuersatz sowie dementsprechend Brutto- und Nettolohn aber gleich sind. Insoweit ergänzt diese Untersuchung die bisherige Forschung durch eine noch nicht in dieser Form untersuchte Fragestellung, eine realitätsnahe Modellierung des Experimentrahmens und eine adäquate Modellierung der Steuerkomplexität.

3. Spezifikation der Forschungsfrage und Experimentdesign

3.1 Spezifikation der Forschungsfrage

Im Rahmen des Experiments werden die Teilnehmer mit einer realitätsnahen Arbeits-Freizeit-Entscheidung konfrontiert. Alle Teilnehmer erhalten während des Experiments den gleichen Bruttolohn für ihre Arbeitsleistung, welcher noch der Besteuerung unterworfen werden muss. Der jeweilige effektive Steuersatz und somit ebenfalls der Nettolohn sind gleich. Nur die Tarifkomplexität des zugrunde liegenden Steuersystems unterscheidet sich in vier verschiedenen Komplexitätsstufen. Die

⁸ *Fochmann et al.* (2010) wählen ein vergleichbares Vorgehen, stellen dabei aber eine andere Fragestellung in den Mittelpunkt.

Komplexität des Steuertarifsystems wird hier durch das Zusammenspiel verschiedener Steuersätze modelliert, um so die Frage zu beantworten, wie die Kombination aus mehreren Steuersätzen (und damit die Steuersatzkomplexität) die Wahrnehmung der Besteuerung bei den Steuerpflichtigen beeinflusst. Folglich stehen bei dieser Untersuchung Steuertarifeffekte im Mittelpunkt.⁹ Bemessungsgrundlageneffekte und Zeiteffekte der Besteuerung werden dabei nicht untersucht.¹⁰ Es wird gerade dieses Vorgehen gewählt, da sich hierfür zahlreiche Beispiele im nationalen¹¹ sowie in internationalen¹² Steuertarifsystemen finden lassen, die nahezu alle Steuerpflichtigen betreffen.

Die Modellierung eines progressiven Tarifverlaufs wird dabei in dieser Untersuchung bewusst ausgeklammert, da es das Ziel dieses Beitrags ist, die steuerlichen Komplexitätswirkungen durch das Zusammenspiel verschiedener Steuersätze zu untersuchen. Dieses praxisnahe Beispiel steuerlicher Tarifkomplexität soll nicht durch Fragen und Probleme der Progression weiter ausgeweitet werden.¹³ Auch im Hinblick auf die kognitiven Fähigkeiten der Teilnehmer erscheint dieses Vorgehen gerechtfertigt, da das Berechnen progressiver Steuersätze eine für diese Fragestellung unverhältnismäßige und auch nicht zielführende Erweiterung der Komplexität darstellen würde.

Es stellt sich folglich die Frage, wie Steuerpflichtige reagieren, wenn sie mit einem komplexen Steuersatzsystem konfrontiert werden. Es ist davon auszugehen, dass mit steigender Komplexität weniger Steuerpflichtige in der Lage sind, ihren effektiven Steuersatz zu berechnen, und folglich mehr Steuerpflichtige auf Heuristiken und persönliche Einschätzungen bzw. Vereinfachungen

⁹ Zur Typisierung steuerlicher Effekte in Zeiteffekte, Bemessungsgrundlageneffekte und Steuertarifeffekte vgl. *Wagner* (1984), S. 211f. sowie *Wagner* (1999), S. 478f.

¹⁰ *Blaufus et al.* (2009) zeigen bereits empirisch, dass Steuertarifeffekte von Steuerpflichtigen meist überschätzt werden, wohingegen Bemessungsgrundlageneffekte eher unterschätzt werden. Auch vor diesem Hintergrund scheint die Modellierung in diesem Experiment über den Steuertarif angemessen, um so die Teilnehmer für die Beachtung der Steuer zu sensibilisieren.

¹¹ Im nationalen Steuertarifsystem ist hierbei das Zusammenwirken von Einkommensteuer und Solidaritätszuschlag sowie Kirchensteuer zu nennen. Auch ein Vergleich mit dem Zusammenspiel von Einkommensteuer und Gewerbesteuer sowie deren Anrechnung nach § 35 EStG liegt dabei nahe. Diese Form der Steuersatzkomplexität wird in politischen Debatten, neben Vereinfachungsvorschlägen, welche allein die Bemessungsgrundlage betreffen, häufig aufgegriffen. So wurde die Selbstabzugsfähigkeit der Gewerbesteuer mit der Unternehmensteuerreform 2008 abgeschafft, was deren Berechnung stark vereinfachte. Weiterhin wird seit langem über eine generelle Abschaffung bzw. Reform der Gewerbesteuer und deren Integration in die Einkommen- bzw. Körperschaftsteuer diskutiert. Vgl. *Maiterth* (2003). Auch das von der FDP vorgeschlagene Stufenmodell zielt zu einem großen Teil auf Steuervereinfachungen durch Vereinfachung des Steuersatzes ab. Vgl. *FDP* (2010), S. 6.

¹² Im internationalen Kontext könnte es sich hier bspw. um das Zusammenspiel zwischen „local taxes“ und „federal taxes“ in den USA handeln, oder aber auch um das Zusammenspiel inländischer und ausländischer Steuersätze bei der Besteuerung internationaler Tätigkeiten, bei welchen es zur Anwendung des Anrechnungs- oder Abzugsverfahrens kommt.

¹³ Durch die Annahme linearer Steuersätze entsprechen sich Grenz-, Durchschnitts- und Differenzsteuersatz. Verwechslungen dieser Steuersätze seitens der Steuerpflichtigen wurden in der Literatur bereits zahlreich untersucht (vgl. Literaturüberblick) und sollen in diesem Beitrag nicht thematisiert und untersucht werden. Aus diesem Grund wird von progressiven Steuertarifen abgesehen.

zurückgreifen. Dabei ist es besonders von Interesse, ob die von den Steuerpflichtigen angewandten Heuristiken eher zu einer Über- oder Unterschätzung der tatsächlichen Steuerbelastung führen, was im Rahmen des Experiments die Arbeits-Freizeit-Entscheidung signifikant beeinflussen sollte.

In einer rationalen Welt, in der die Steuerpflichtigen die zu leistenden Steuerzahlungen richtig kalkulieren und antizipieren, sollte es keine signifikanten Unterschiede zwischen den verschiedenen Gruppen in diesem Experiment hinsichtlich des Arbeitsangebotes geben. Da die Berechnung des Nettolohns aber abhängig vom individuellen Steuertarifsystem mit unterschiedlichem Aufwand der Steuerberechnung verbunden ist, ist zu erwarten, dass sich insbesondere die Teilnehmer, welche mit einem komplexen Steuertarifsystem konfrontiert sind, eher eine Heuristik oder Vereinfachungsregel anwenden. Diese Teilnehmer werden dann ihr Arbeitsangebot anhand eines aus ihrer Sicht unsicheren Nettolohns wählen.

Die Forschungsfrage lässt sich somit weiter spezifizieren, indem untersucht wird, ob dieses Vorgehen eher zu einer Ausweitung oder Einschränkung des Arbeitsangebotes führt, was aus einer signifikanten Über- oder Unterschätzung des Steuersatzes und damit des Nettolohns resultieren kann.¹⁴ Die Aufdeckung der Richtung sowie die Quantifizierung dieses Verhaltens ist das Ziel dieses Beitrags. Außerdem werden Anhaltspunkte für die Grundlagen der individuell angewandten Heuristik herausgearbeitet. So könnte z. B. ein Ankereffekt¹⁵ vorliegen, der dazu führt, dass sich die Teilnehmer aus Vereinfachungsgründen an dem höchsten oder zuerst genannten Steuersatz orientieren und die sonstigen Steuern vernachlässigen. Auch eine vereinfachte Orientierung an der Summe der einzelnen Steuersätze wäre möglich. Bei der Modellierung der Steuertarifsysteme wurde deshalb

¹⁴ Die Wirkungsrichtung, die eine Über- oder Unterschätzung des Steuersatzes auf das Arbeitsangebot hat, kann im Vorfeld nicht eindeutig bestimmt werden. Bei den Teilnehmern kann eine Fehleinschätzung der zu zahlenden Steuerlast und damit des Nettolohns sowohl einen Einkommens- wie auch einen Substitutionseffekt auslösen. Wird bspw. der Steuersatz zu hoch und damit der Nettolohn zu niedrig eingeschätzt, kann der Einkommenseffekt zu einer Verringerung (da sich die Arbeit weniger lohnt) oder Ausweitung des Arbeitsangebotes (bspw. wenn ein bestimmter Mindestnettoloohn erzielt werden soll) führen. Der Substitutionseffekt führt dagegen immer zu einer Einschränkung des Arbeitsangebotes, da mit sinkendem Lohn Arbeitszeit durch Freizeit ersetzt wird. Ob der Einkommenseffekt und der Substitutionseffekt in die gleiche Richtung wirken oder ob der Substitutionseffekt vom Einkommenseffekt (über-)kompensiert wird hängt vom individuellen Verhalten der Teilnehmer ab. Zur Theorie vgl. bereits *Hicks* (1935). Empirische Überprüfungen der Arbeitsangebotsreaktion auf Lohnänderungen finden sich bspw. bei *Break* (1957), *Hausman* (1981), *Killingsworth* (1983), *Kimmel* und *Kniesner* (1998) sowie *Ziliak* und *Kniesner* (1999). Ein Überblick findet sich bei *Keane* (2010). Insgesamt deuten die bisherigen Untersuchungen eher auf einen positiven Zusammenhang zwischen Nettolohn und Arbeitsangebot hin, so dass auch hier die Vermutung einer Arbeitsangebotsreduktion (-ausweitung) bei (Unter-) Überschätzung der Steuerbelastung und somit (Über-) Unterschätzung des Nettolohns als Ausgangshypothese übernommen werden kann.

¹⁵ Ein Ankereffekt (*Tversky* und *Kahnemann* (1974)) bedeutet, dass sich die Probanden in der unsicheren Situation einen Startpunkt suchen, welchen diese dann beim Treffen ihrer Entscheidung zu stark gewichten. In dieser Untersuchung könnte der erste und höchste Steuersatz (Einkommensteuersatz) einen solchen Anker darstellen. Weitere empirische Überprüfungen eines Ankereffekts bei verschiedenen Fragestellungen finden sich bspw. bei *McCaffery* und *Baron* (2003), *Epley* und *Gilovich* (2006) sowie bei *Blaufus et al.* (2009).

darauf geachtet, dass ein solches Vorgehen der Teilnehmer untersucht werden kann. Der am Ende des Experiments vorgelegte Fragebogen enthält weiterhin spezifische Fragen, die Aufschluss darüber geben, inwieweit das Steuertarifsystem überhaupt Beachtung bei der Arbeitsangebotsentscheidung gefunden hat und ob der effektive Steuersatz vor Beginn der Arbeitstätigkeit berechnet wurde.

3.2 Experimentdesign

Das Experiment wurde in fünf Durchgängen von April 2011 bis Juni 2011 in den Räumen der Freien Universität Berlin durchgeführt. Dabei nahmen insgesamt 96 Probanden an dem Experiment teil. Bei den Probanden handelt es sich überwiegend um Studenten im zweiten Semester des Bachelorstudiengangs Betriebswirtschaftslehre.¹⁶ Auch Studenten anderer Fachbereiche waren unter den Teilnehmern vertreten. Da keine signifikanten Unterschiede zwischen den einzelnen Durchläufen sowie zwischen den Studenten der Betriebswirtschaftslehre und den Studenten anderer Fachbereiche existieren, werden diese Gruppen im folgenden zusammengefasst.¹⁷

Das Experiment wurde in den PC-Pools des Fachbereichs Wirtschaftswissenschaft durchgeführt. Die Teilnehmer wurden nach dem Zufallsprinzip auf die zur Verfügung stehenden Personal-Computer verteilt.¹⁸ Die jeweiligen Arbeitsplätze der Teilnehmer waren durch einen Sichtschutz voneinander abgetrennt, außerdem war die Kommunikation zwischen den Teilnehmern für die gesamte Dauer des Experiments untersagt. Bei keinem der Durchläufe waren mehr als 25 Teilnehmer gleichzeitig anwesend. Es standen bei jedem Durchlauf mehrere Experimentleiter zur Verfügung, welche den Ablauf des Experiments überwachten und auftretende Fragen beantworteten.

Zu Beginn des Experiments erhielten die Teilnehmer eine Anleitung ausgehändigt, die alle notwendigen Informationen zum Ablauf des Experiments beinhaltet.¹⁹ Im Anschluss daran wurde den Teilnehmern ausreichend Zeit gegeben, diese Anleitung sorgfältig durchzulesen und auftretende

¹⁶ Der Kritikpunkt, dass Studenten als Experimentteilnehmer ungeeignet bzw. nicht repräsentativ sind, wird in der Literatur bestritten. Vgl. *Plott* (1987), *Alm* (2010).

¹⁷ Zur Überprüfung wurde die Verteilung relevanter Variablen zwischen den einzelnen Durchläufen sowie den entsprechenden Studiengängen der Teilnehmer mit Hilfe von Kruskal-Wallis-Tests und Mann-Whitney-U-Tests überprüft. Es konnten keinerlei signifikante Unterschiede zwischen den Kategorien festgestellt werden. Des Weiteren wurde die Zugehörigkeit zum Studiengang Betriebswirtschaftslehre oder einem sonstigen Studiengang in die durchgeführte Regressionsanalyse aufgenommen, wobei sich allerdings auch dort kein signifikanter Einfluss nachweisen lässt.

¹⁸ Es gab keine signifikanten Unterschiede bei dieser Verteilung hinsichtlich der demografischen Merkmale: Alter, Geschlecht, Einkommen und Schulbildung bzw. Studium.

¹⁹ Die Anleitung findet sich im Anhang.

Fragen zu stellen. Die Teilnehmer wurden mit einer realen Arbeits-Freizeit-Entscheidung konfrontiert. Sie bekamen die Aufgabe, Lösungsbögen einer Multiple-Choice-Klausur zu Korrekturzwecken in eine vorgefertigte Exceldatei zu übernehmen. Jeder Lösungsbogen bestand aus 40 Fragen mit jeweils 6 Antwortmöglichkeiten, von denen jeweils eine Antwortmöglichkeit angekreuzt war. Bei der Wahl dieser Form einer Arbeitsaufgabe war es zum einen wichtig, dass die Teilnehmer das Gefühl hatten, eine realitätsnahe Aufgabe auszuführen und zum anderen sollte eine Aufgabe gewählt werden, welche keine Vorkenntnisse oder besondere Fähigkeiten erfordert. Für jeden fehlerfrei digitalisierten Bogen erhielten die Teilnehmer einen Bruttolohn in Höhe von 0,30 Euro, was ihnen ebenfalls in der Anleitung mitgeteilt wurde. Als Freizeitalternative zu dieser Arbeitsaufgabe wurden die Teilnehmer darüber informiert, dass diese das Experiment jederzeit beenden und nach Hause gehen können. Eine Zeitbeschränkung gab es nicht, jedem Teilnehmer stand es somit frei, so lange und soviel zu arbeiten wie er wollte.²⁰ So konnte eine realitätsnahe Entscheidungssituation geschaffen werden. Bei jedem Durchgang waren speziell instruierte zusätzliche Teilnehmer anwesend, die die Aufgabe hatten, nach 30 Minuten das Experiment zu beenden. Dadurch sollte die ggf. bestehende Hürde zur frühzeitigen Beendigung des Experiments beseitigt werden.

Der Bruttoarbeitslohn musste von den Teilnehmern versteuert werden. Dazu wurde den Teilnehmern im Anschluss an die Instruktionen noch eine Steuersystemkarte ausgehändigt, welche das für sie geltende Steuertarifsystem beinhaltet. Eine Auskunft über den Nettolohn selbst wurde den Teilnehmern nicht gegeben. Die Steuersystemkarten wurden jeweils zufällig auf die Teilnehmer verteilt, wobei darauf geachtet wurde, dass bei jedem Durchlauf, soweit möglich, jedem Steuersystem gleich viele Teilnehmer zugeordnet wurden.²¹ Dadurch konnte im Anschluss an das Experiment überprüft werden, ob mögliche Effekte auf durchlaufspezifische Unterschiede zurückzuführen sind.

Es gab insgesamt vier verschiedene Steuersystemkarten. Dabei unterschieden sich die Steuersysteme hinsichtlich ihrer Komplexität in der Form, dass jeweils eine unterschiedliche Anzahl an Steuersätzen zu beachten waren. Diese einzelnen Steuersätze mussten zur Berechnung des Gesamtsteuersatzes teilweise additiv oder multiplikativ verknüpft werden. Bei einer der Varianten (Steuersystem D) war eine Steuerzahlung zudem noch abzugsfähig von der Bemessungsgrundlage einer

²⁰ Das Experiment wurde ohne vorherige Ankündigung jeweils nach ca. 3 Stunden beendet, was über alle Durchläufe nur 5 Teilnehmer betraf. Alle anderen Teilnehmer haben das Experiment schon vorher, ohne Kenntnis des geplanten Abbruchzeitpunkts, beendet.

²¹ Die Steuertarifsystemkarten finden sich im Anhang.

weiteren Steuer. Die Besonderheit bei diesem Vorgehen bestand darin, dass der effektive Steuersatz bei allen Varianten gleich war und konstant 60% betrug.²² Lediglich die Darstellung bzw. der Komplexitätsgrad variierte zwischen den Steuertarifsystemen, was eine Berechnung des effektiven Steuersatzes an einen unterschiedlichen Aufwand bzw. unterschiedliche Kenntnisse knüpfte.

Variante	Einkommensteuer	Zusatzsteuer auf das Einkommen	Zuschlagsteuer auf die Einkommensteuer	Landessteuer auf das Einkommen	Gemeindesteuer mit Abzugsfähigkeit bei der Einkommensteuer	Effektiver Steuersatz	Komplexitätsmodellierung
A	60%	---	---	---	---	60,0%	Einfacher Steuersatz
B	40%	20%	---	---	---	60,0%	Additive Verknüpfung
C	40%	---	20%	12%	---	60,0%	Additive und multiplikative Verknüpfung
D	20%	---	20%	20%	20%	59,2%	Additive und multiplikative Verknüpfung sowie Abzugsfähigkeit

Tabelle 1 - Steuertarifsysteme

Während bei Variante A lediglich ein pauschaler Steuersatz in Höhe von 60% erhoben wurde, musste dieser bei Variante B über eine einfache additive Verknüpfung ermittelt werden ($40\% + 20\% = 60\%$). Für die Ermittlung des effektiven Gesamtsteuersatzes war bei Variante C dann zusätzlich eine multiplikative Rechenoperation notwendig. Hier ergab sich der Gesamtsteuersatz als $40\% \cdot (1 + 20\%) + 12\% = 60\%$. Bei Variante D wurde dann das Rechenverfahren nochmals in der Form erweitert, dass zusätzlich die Abzugsfähigkeit der Gemeindesteuer bei der Einkommensteuer beachtet werden musste. Die effektive Gesamtsteuerbelastung ergab sich somit als:

²² Der effektive Steuersatz wurde bewusst in dieser Höhe gewählt, um die Teilnehmer implizit für die Beachtung der Steuerzahlung zu sensibilisieren.

$20\% \cdot (1 + 20\%) + 20\% + 20\% - 20\% \cdot 20\% \cdot (1 + 20\%) = 59,2\% \approx 60\%$ ²³. Die Komplexität des Steuertarifsystems nimmt folglich – bei praktisch gleicher Gesamtbelastung – von Variante A nach Variante D zu, wobei die Komplexität hier durch die Anzahl der zu berücksichtigenden Steuersätze sowie deren Interaktionen modelliert wurde. Dabei wurden die notwendigen Rechenschritte die zur Berechnung der effektiven Gesamtsteuerbelastung notwendig sind immer weiter ausgebaut und um Rechenschritte, wie sie auch im deutschen Steuersystem bei der Interaktion verschiedener Steuersätze zu finden sind, ergänzt. Dieses Vorgehen erlaubt folglich eine stufenweise Komplexitätsmodellierung. Bei den Steuersystemkarten selbst wurde versucht, die sprachlichen Formulierungen so einfach wie möglich zu halten und steuerrechtliche Fachtermini soweit möglich zu vermeiden.²⁴ Es wurde bewusst auf die Angabe komplexer Steuertariffunktionen, bspw. unter Beachtung von Progressionswirkungen, verzichtet, da dies für die Beantwortung der Forschungsfrage, aufgrund einer unverhältnismäßigen Erhöhung der Komplexität, kontraproduktiv wäre.

Nachdem die Teilnehmer das für sie geltende Steuertarifsystem erhalten hatten, wurden jedem Teilnehmer Lösungsbögen der Klausur in ausreichender Stückzahl übergeben und das Experiment wurde gemeinsam begonnen.²⁵ Die Startzeit wurde vom Experimentleiter zur späteren Auswertung der jeweiligen Arbeitszeit notiert. Jeder Teilnehmer wurde nach seiner individuellen Beendigung der Arbeitstätigkeit in einen separaten Raum geführt, wo dieser entsprechend seiner Arbeitsleistung ausbezahlt wurde und einen zusätzlichen Fragebogen bekam. Der Fragebogen beinhaltete sowohl demografische Fragen wie auch Fragen zur Zufriedenheit mit der Arbeitstätigkeit und dem entsprechenden Lohn und Steuertarifsystem. Alle Teilnehmer stimmten der Beantwortung des Fragebogens zu.²⁶ Die individuelle Arbeitszeit sowie die Arbeitsleistung (gemessen anhand der (fehlerfrei) digitalisierten Bögen) wurde exakt protokolliert.

²³ Der Rundungsfehler bei dieser Variante beträgt 0,8 Prozentpunkte und kann damit vernachlässigt werden.

²⁴ Auch die Benennung der verschiedenen Steuerarten (Einkommensteuer, Zusatzsteuer, Zuschlagsteuer, Landessteuer, Gemeindesteuer) wurde so allgemein wie möglich gehalten, um einen potenziellen Labeling- oder Affektationseffekt einer Steuer zu vermeiden. Vgl. dazu *Hundsdoerfer et al. (2010)*.

²⁵ Die Lösungsbögen waren so konzipiert, dass keine Systematik erkennbar war, die die Arbeitsaufgabe vereinfacht hätte. Sonstige Hilfsmittel waren nicht zugelassen.

²⁶ Der Fragebogen befindet sich im Anhang.

4. Ergebnisse und Diskussion

Als Maßstab für die Messung der erbrachten Arbeitsleistung kann die Arbeitszeit oder die Anzahl der digitalisierten Bögen herangezogen werden. Da die zu bewältigende Arbeitsaufgabe selbst zwar keine besonderen Fähigkeiten voraussetzte, aber dennoch ein hohes Maß an Konzentration erforderte, ist weiterhin die Produktivität und die Fehlerhäufigkeit von Interesse. Da die geleistete Arbeit vor Auszahlung des entsprechenden Arbeitsentgeltes auf Fehlerfreiheit geprüft wurde, wurden dementsprechend auch nur die fehlerfrei digitalisierten Bögen vergütet. Die folgende Tabelle stellt die gesamte erbrachte Arbeitsleistung (Anzahl der digitalisierten Bögen) sowie die entsprechende Arbeitszeit während des Experiments dar:

	Mittelwert	Median	Minimum	Maximum	Standardabweichung	Summe
Arbeitsminuten	102,84	102,50	11,00	181,00	47,27	9.873,00
Digitalisierte Bögen	74,97	66,50	4,00	199,00	48,09	7.197,00
Fehlerhafte Bögen	11,47	10,50	0,00	45,00	7,52	1.101,00
Bruttolohn	19,05	17,10	0,30	52,80	13,09	1.828,80
Nettolohn	7,62	6,84	0,12	21,12	5,24	731,52

Tabelle 2 - Gesamte Arbeitsleistung im Experiment

Die durchschnittliche Produktivität über alle Teilnehmer beträgt 0,686 Bögen pro Minute, wobei diese auf einer Intervallbreite von 0,221 Bögen pro Minute bis 1,725 Bögen pro Minute liegt und somit stark zwischen den Teilnehmern variiert. Es kann davon ausgegangen werden, dass die Produktivität bei dieser Aufgabe insbesondere von motorischen Fähigkeiten, bspw. der Hand-Augen-Koordination, abhängt. Weiterhin ist es für die Auswertung der Ergebnisse von Interesse, wie sich die Produktivität im Zeitverlauf verhält. Wenn sich die Produktivität im Zeitablauf stark ändert, würde das bedeuten, dass sich die beiden Messgrößen für das Arbeitsangebot (Arbeitszeit und Arbeitsleistung) zwischen Teilnehmern mit hohem Arbeitsangebot und Teilnehmern mit niedrigem Arbeitsangebot nicht linear verhalten. So würde ein Teilnehmer im Vergleich zu einem anderen bei doppelter Arbeitszeit nicht genau das doppelte an Arbeitsleistung aufweisen, selbst wenn die beiden Personen vom Grundsatz her mit der gleichen Produktivität starten würden. Dies könnte die Güte der Arbeitsangebotsmessung beeinflussen und muss somit berücksichtigt werden.

Um einen solchen Effekt zu überprüfen, wurden die Probanden entsprechend der gearbeiteten Zeit in sechs Gruppen eingeteilt. Um in jedem Zeitintervall ausreichend Teilnehmer für diese Darstellung zu haben, wurden die Zeitintervalle jeweils in 30-Minuten-Schritte eingeteilt. So kann eine Aussage darüber getroffen werden, inwieweit die erbrachte Arbeitsleistung auch von nicht steuerinduzierten bzw. nicht monetären Effekten abhängt. Die folgende Tabelle und Abbildung stellt diese Ergebnisse dar. Dabei bezieht sich die Bruttonproduktivität auf alle digitalisierten Bögen, die Nettoproduktivität hingegen nur auf die fehlerfreien Bögen.

Zeitintervall (in Minuten)	N	Bruttonproduktivität (Nettoproduktivität) in Bögen pro Minute				
		Mittelwert	Median	Minimum	Maximum	Standardabweichung
00 - 29	6	0,454 (0,261)	0,405 (0,261)	0,304 (0,044)	0,821 (0,500)	0,188 (0,162)
30 - 59	14	0,543 (0,375)	0,474 (0,355)	0,370 (0,140)	0,816 (0,667)	0,142 (0,146)
60 - 89	22	0,649 (0,537)	0,689 (0,562)	0,221 (0,221)	0,943 (0,725)	0,198 (0,141)
90 - 119	14	0,738 (0,641)	0,645 (0,546)	0,472 (0,366)	1,725 (1,560)	0,333 (0,320)
120 - 149	20	0,806 (0,684)	0,772 (0,653)	0,333 (0,225)	1,363 (1,181)	0,245 (0,224)
150 +	20	0,741 (0,648)	0,719 (0,633)	0,408 (0,284)	1,167 (1,024)	0,207 (0,200)
Gesamt	96	0,686 (0,565)	0,682 (0,557)	0,221 (0,044)	1,725 (1,560)	0,245 (0,240)

Tabelle 3 - Produktivitäten im Zeitablauf

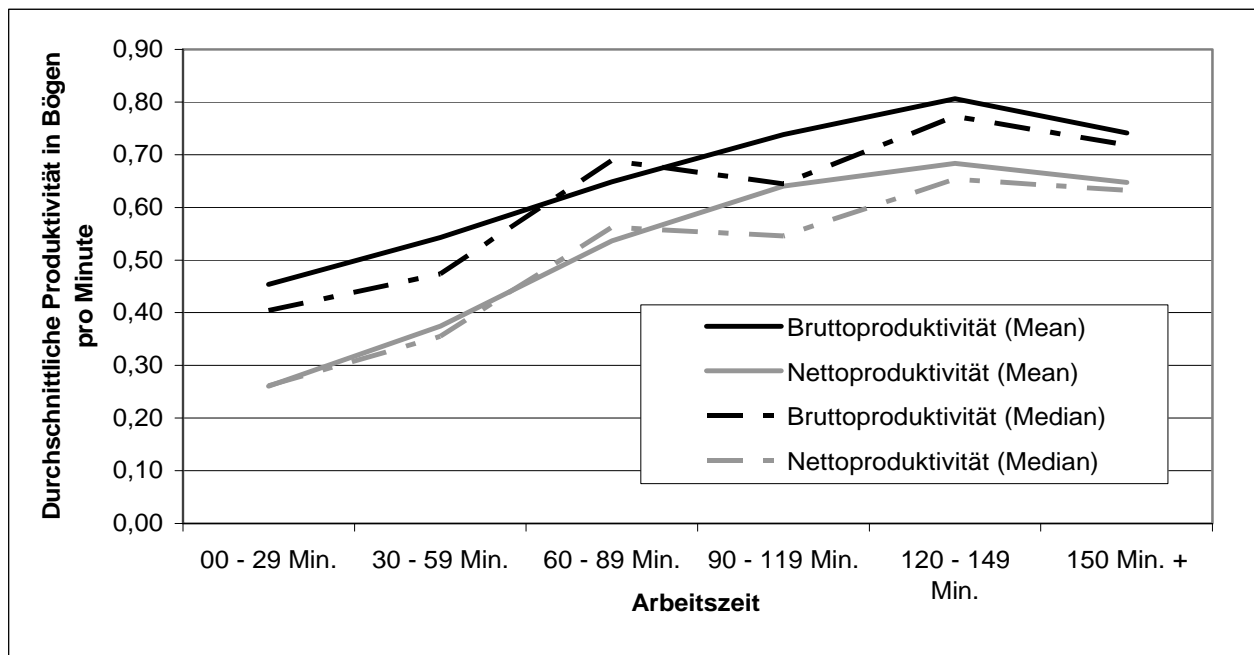


Abbildung 1 - Produktivitäten im Zeitablauf

Es wird deutlich, dass die Teilnehmer, die eine lange Zeit gearbeitet haben, im Mittel eine höhere Produktivität aufweisen als die Teilnehmer, die das Experiment schon nach kurzer Zeit beendet haben. Dieses Ergebnis könnte auf einen positiven Lerneffekt bei der Bearbeitung der Aufgabe hinweisen.²⁷ Bei den Teilnehmern mit der längsten Arbeitsdauer nimmt die durchschnittliche Produktivität dann wieder leicht ab. Ursächlich dafür können vor allem Ermüdungserscheinungen sein, da die Aufgabe ein hohes Maß an Konzentration erforderte und die Teilnehmer so nach einem anfänglichen Produktivitätsanstieg wieder einen Abfall der Produktivität aufweisen. Diese Vermutung deckt sich auch mit den Angaben der Teilnehmer im anschließend ausgefüllten Fragebogen, in dem zahlreiche Teilnehmer über Konzentrationsschwierigkeiten, Kopfschmerzen oder auftretende Verspannungen berichteten, welche nach einer langen Arbeitszeit mit zur Beendigung der Teilnahme beigetragen haben. Dieser Produktivitätseffekt wird im weiteren Verlauf der Untersuchung somit zu berücksichtigen sein.

Das zentrale Ergebnis zur Beantwortung der hier untersuchten Fragestellung ist, dass die Teilnehmer an dem Experiment länger gearbeitet und mehr Bögen digitalisiert haben, je einfacher das Steuertarifsystem hinsichtlich des Zusammenspiels verschiedener Steuersätze war. Im Mittel wurde

²⁷ Ob es sich wirklich um einen Lerneffekt handelt kann nur vermutet werden, da die Daten keinen Aufschluss darüber geben, wie sich die Produktivität der einzelnen Teilnehmer im Zeitablauf entwickelt hat. Denkbar wäre hier auch, dass die Teilnehmer, die von Anfang an eine höhere Produktivität aufweisen, ebenfalls eine höhere Arbeitsbereitschaft haben.

bei dem einfachsten Steuertarifsystem (Steuersystem A) die meiste Zeit mit der Arbeitsaufgabe verbracht und die meisten Bögen digitalisiert. Bei dem komplexen Steuertarifsystem D fällt hingegen auf, dass im Mittel weniger gearbeitet wurde.

	Mittelwert	Median	Standardabweichung	Minimum	Maximum
Arbeitsminuten	125,67	138,00	39,62	35,00	181,00
Digitalisierte Bögen	92,79	80,50	41,83	16,00	196,00
A Fehlerhafte Bögen	13,63	13,50	5,78	2,00	24,00
Bruttolohn	23,75	21,45	11,62	4,20	51,60
Nettolohn	9,50	8,58	4,65	1,68	20,64
Arbeitsminuten	117,54	113,50	38,23	54,00	171,00
Digitalisierte Bögen	80,08	74,00	45,05	20,00	198,00
B Fehlerhafte Bögen	11,54	10,50	6,41	2,00	26,00
Bruttolohn	20,56	18,45	12,51	2,40	52,80
Nettolohn	8,23	7,38	5,00	0,96	21,12
Arbeitsminuten	97,67	92,50	50,42	15,00	174,00
Digitalisierte Bögen	74,88	61,00	52,41	6,00	199,00
C Fehlerhafte Bögen	11,33	8,00	10,12	2,00	45,00
Bruttolohn	19,06	15,30	13,93	0,30	50,70
Nettolohn	7,63	6,12	5,57	0,12	20,28
Arbeitsminuten	70,50	66,00	42,05	11,00	139,00
Digitalisierte Bögen	52,13	28,50	46,09	4,00	188,00
D Fehlerhafte Bögen	9,38	8,00	6,86	0,00	23,00
Bruttolohn	12,83	6,90	12,50	0,90	51,00
Nettolohn	5,13	2,76	5,00	0,36	20,40

Tabelle 4 - Arbeitsleistung nach Gruppen

Die folgenden Abbildungen sollen die Verteilung der wesentlichen Messgrößen des Experiments in Abhängigkeit von dem jeweiligen Steuertarifsystem nochmals verdeutlichen:

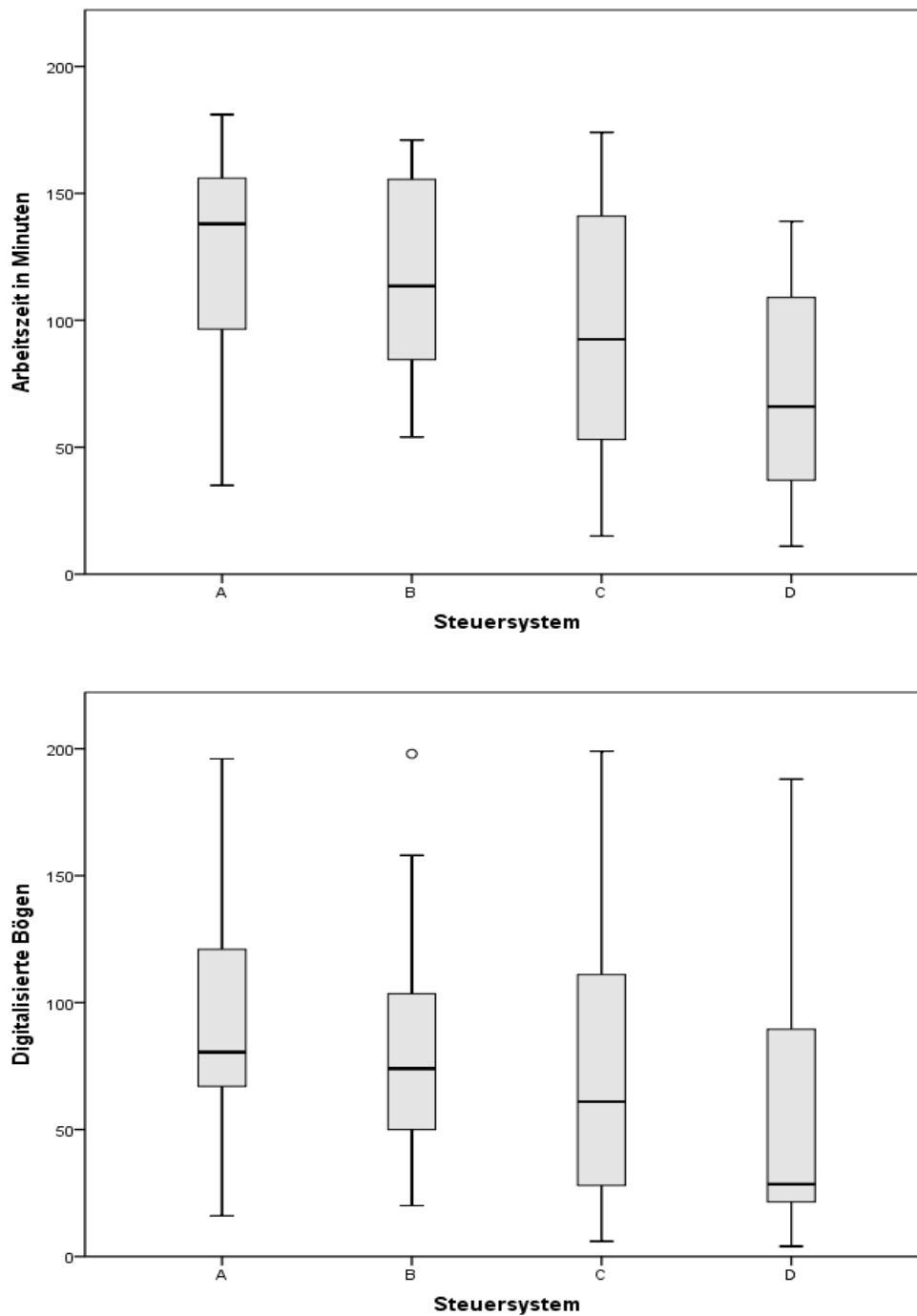


Abbildung 2 - Arbeitszeit und Arbeitsleistung

Um zu überprüfen, inwieweit die Unterschiede zwischen der Arbeitsbereitschaft unter den verschiedenen Steuertarifsystemen statistisch signifikant sind, wurden paarweise Mann-Whitney-U-

Tests²⁸ durchgeführt. Die folgende Tabelle gibt einen Überblick über die Ergebnisse. Dabei wurde sowohl auf einen Unterschied hinsichtlich der Arbeitszeit wie auch der Arbeitsleistung (Anzahl der digitalisierten Bögen) und der Produktivität getestet.²⁹ Die Tabelle enthält jeweils die p-Werte der Testverfahren.

Verglichene Gruppen	Arbeitszeit	Arbeitsleistung	Produktivität
A - B	0,470	0,265	0,180
A - C	0,051*	0,101	0,529
A - D	0,000***	0,001***	0,248
B - C	0,146	0,585	0,421
B - D	0,000***	0,015**	0,942
C - D	0,058**	0,083*	0,483

***Signifikant auf dem 1% Niveau, **5% Niveau, *10% Niveau

Tabelle 5 - Paarweise Gruppenvergleiche

Es wird deutlich, dass sowohl die Arbeitszeit als auch die Arbeitsleistung zwischen einigen Gruppen signifikant verschieden ist.³⁰ Während zwischen dem einfachen Steuertarifsystem (Steuersystem A) und dem um eine additive Operation erweiterten Steuertarifsystem (Steuersystem B) kein signifikanter Unterschied festzustellen ist, wird insbesondere die signifikant geringere Arbeitszeit und Arbeitsleistung bei den Teilnehmern unter Steuertarifsystem D deutlich. Alle anderen Steuertarifsystemgruppen unterscheiden sich signifikant von Steuertarifsystem D, welches den höchsten Komplexitätsgrad aufwies. Zwischen den Steuertarifsystemen B und C lässt sich des Weiteren kein signifikanter Unterschied feststellen, wobei sich diese Steuertarifsysteme durch die Einführung einer weiteren Zuschlagsteuer unterscheiden. Da sich das Arbeitsangebot zwischen Gruppe A und Gruppe B gar nicht und zwischen Gruppe A und Gruppe C nur schwach signifikant unterscheidet,

²⁸ Bei dem Mann-Whitney-U-Test handelt es sich um ein parameterfreies statistisches Testverfahren. Dieses Testverfahren überprüft, ob zwei unabhängige Verteilungen zu derselben Grundgesamtheit gehören. Aufgrund der Beschaffenheit der Daten ist die Anwendung eines parameterfreien Testverfahrens geboten. Vgl. dazu auch *Mann* und *Whitney* (1947) sowie *Wilcoxon* (1945).

²⁹ Die Produktivität dient dabei als relativer Maßstab. Damit kann untersucht werden, ob sich Gruppenunterschiede hinsichtlich Arbeitszeit und Arbeitsleistung linear zwischen den Gruppen entwickeln.

³⁰ Es wurde auch die Fehlerquote auf einen Steuertarifsystembedingten Unterschied überprüft. Hier konnten aber keine signifikanten Abweichungen identifiziert werden.

kann ein möglicher Ankereffekt des höchsten und zuerst genannten Steuersatzes (Einkommensatz) anhand dieser Ergebnisse nicht bestätigt werden.³¹

Hinsichtlich der Produktivität konnten keine Unterschiede zwischen den Steuertarifsystemgruppen gefunden werden. Es deutet nichts darauf hin, dass das Steuertarifsystem auch einen Einfluss auf die Schnelligkeit bei der Arbeitsaufgabe hat. Dies scheint der Tatsache geschuldet, dass die Produktivität vor allem von physiologischen Ursachen restringiert wird und diese von den Teilnehmern nur bedingt beeinflusst werden können. Lediglich die Arbeitszeit scheint einen messbaren Einfluss auf die Produktivitätsrate der Teilnehmer auszuüben. Die folgende Tabelle stellt die Produktivitäten nochmals getrennt nach Zeit und Steuertarifsystem dar.

Zeitintervall ³²	A	B	C	D
0 - 59 Min.	0,457	0,405	0,483	0,560
60 - 119 Min.	0,722	0,611	0,684	0,731
120 + Min.	0,755	0,728	0,857	0,788

Tabelle 6 - Produktivitäten nach Zeitintervall und Steuertarifsystem

Aus der Tabelle wird ersichtlich, dass es keine auffallenden Produktivitätsunterschiede zwischen den einzelnen Gruppen gibt. Allerdings sollte berücksichtigt werden, dass die mittlere Produktivität bei Teilnehmern mit längerer Arbeitszeit höher ist als bei Teilnehmern mit kurzer Arbeitszeit. Um diesem Umstand innerhalb der Analyse Rechnung zu tragen, wurden die oben durchgeführten Gruppenvergleiche nochmals mit einer um zeitliche Produktivitätseffekte bereinigten Arbeitszeit bzw. Arbeitsleistung durchgeführt.³³ Dazu wurde die mittlere Produktivität über alle Teilnehmer berechnet und dann die Arbeitszeit bzw. -leistung entsprechend dieser Durchschnittsproduktivität korrigiert. Dabei wurde folglich zum einen für jeden Teilnehmer die Arbeitszeit berechnet, die er bei durchschnittlicher Produktivität gebraucht hätte, um die gleiche Anzahl an Bögen zu digitalisieren, die er tatsächlich geschafft hat. Zum anderen wurde die Anzahl der Bögen ent-

³¹ Folgt man der in der überwiegenden Literatur vertretenen Meinung, dass eine Erhöhung des Nettolohns zu einer Ausweitung des Arbeitsangebots führt, so müsste ein solcher Ankereffekt ebenfalls zu einem höheren Arbeitsangebot bei Gruppe D im Vergleich zu allen anderen Gruppen führen. Auch danach kann ein Ankereffekt somit nicht bestätigt werden.

³² Aufgrund der Teilnehmerzahl wurden die Zeitintervalle bei dieser Gruppenbildung größer gefasst.

³³ Dieses Vorgehen stellt eine Normierung der mutmaßlich durch exogene Einflüsse getriebenen Produktivitätsrate dar und dient der Überprüfung der Robustheit der gefundenen Ergebnisse.

sprechend der tatsächlichen Arbeitszeit bei durchschnittlicher Produktivität angepasst. Dieses Vorgehen soll mögliche Messverzerrungen durch Produktivitätseffekte ausschließen.

Die korrigierte Arbeitszeit für Teilnehmer i ergibt sich dann als:

$$\text{Arbeitszeit}_i = \frac{\text{Bögen}_i}{\varnothing \text{ Produktivität}}$$

und die korrigierte Anzahl an Arbeitsleistung als:

$$\text{Bögen}_i = \text{Arbeitszeit}_i \cdot \varnothing \text{ Produktivität} .$$

Die folgende Tabelle stellt die mittlere Arbeitszeit sowie mittlere Arbeitsleistung in der korrigierten Fassung dar.

	Mittelwert	Median	Standardabweichung	Minimum	Maximum	
A	Arbeitsminuten	135,20	117,29	70,95	23,31	285,58
	Digitalisierte Bögen	86,25	94,71	27,19	24,02	124,22
B	Arbeitsminuten	116,68	107,82	65,64	29,14	288,49
	Digitalisierte Bögen	80,67	77,90	26,24	37,06	117,36
C	Arbeitsminuten	109,10	88,88	76,37	8,74	289,95
	Digitalisierte Bögen	67,03	63,49	34,60	10,29	119,42
D	Arbeitsminuten	75,95	41,53	41,53	5,83	273,92
	Digitalisierte Bögen	48,39	45,40	45,30	7,55	95,40

Tabelle 7 - Arbeitszeit und Arbeitsleistung in korrigierter Form

Diese Werte stellen die notwendige Arbeitszeit bzw. die vollbrachte Arbeitsleistung unter der Bedingung dar, dass alle Teilnehmer mit der gleichen Produktivität gearbeitet hätten und es keine exogenen sowie endogenen Einflüsse auf die Produktivität einzelner Teilnehmer gäbe. In diesem Fall wird unterstellt, dass die Teilnehmer ihr Arbeitsangebot entweder ausschließlich an einer festen Zeit ausgerichtet haben oder aber sich eine bestimmte Stückzahl bei der Arbeitsleistung als Ziel gesetzt haben. Effekte durch das schnellere oder langsamere Arbeiten werden dabei ausgeblendet. Es sei angemerkt, dass damit auch Produktivitätsunterschiede durch steuerinduzierte Anreize ausgeblendet werden.

Während des Experiments war weder die Kommunikation mit anderen Teilnehmern noch eine alternative Tätigkeit erlaubt, so dass es unwahrscheinlich ist, dass die Produktivität bewusst beeinflusst wurde. Es kann vermutet werden, dass die Teilnehmer während ihrer Teilnahme auch ihre volle Konzentration auf die Arbeitsaufgabe verwendet haben oder aber das Experiment abgebrochen haben. Da die Produktivitätsunterschiede somit eher durch nicht beeinflussbare Restriktionen bestimmt waren, scheint dieses Vorgehen zur Kontrolle der obigen Ergebnisse angemessen.

Hier wurden auch wieder paarweise Gruppenvergleiche mit Hilfe des Mann-Whitney-U-Tests durchgeführt. Die Ergebnisse unterscheiden sich nur schwach im Vergleich zu den vorangegangenen Tests mit der unkorrigierten Messgröße, was wiederum die Ergebnisse weiter bestätigt.

Verglichene Gruppen:	Arbeitszeit	Arbeitsleistung
A - B	0,265	0,470
A - C	0,101	0,051**
A - D	0,001***	0,000***
B - C	0,585	0,146
B - D	0,015**	0,000***
C - D	0,083*	0,058*

***Signifikant auf dem 1% Niveau, **5% Niveau, *10% Niveau

Tabelle 8 - Paarweise Gruppenvergleiche in korrigierter Form

Aus diesen Ergebnissen lässt sich ableiten, dass eine Erhöhung der Komplexität des Steuertarifsystems zur Folge hat, dass die Steuerpflichtigen ihr Arbeitsangebot reduzieren. Obwohl der effektive Steuersatz sowie der Brutto- und Nettolohn in allen vier Steuertarifsystemvarianten gleich ist, zeigt sich ein signifikanter Unterschied. Das Steuertarifsystem D bestand aus vier verschiedenen Steuersätzen, welche durch additive und multiplikative Verknüpfungen sowie die Abzugsfähigkeit einer Steuer bei der Berechnung einer anderen gekennzeichnet war. Dieses Geflecht aus verschiedenen Steuersätzen konnte nur mit großem Aufwand zu einem effektiven Gesamtsteuersatz zusammengefügt werden. Wenn die Teilnehmer dazu nicht in der Lage waren, bzw. ihnen diese Aufgabe als zu mühselig erschien, standen diese vor der Wahl, die Steuern entweder ganz zu vernachlässigen oder eine Vereinfachungsregel anzuwenden bzw. ihr Arbeitsangebot an einer Heuristik auszurichten. Da das Arbeitsangebot in dieser Steuertarifsystemgruppe signifikant unter

dem Angebot in den anderen Gruppen liegt, kann gefolgert werden, dass die Teilnehmer aufgrund der Unsicherheit über ihre tatsächliche Steuerzahlung und somit über ihren Nettolohn eher risikoavers reagieren und den Steuersatz als zu hoch und die Arbeit damit als weniger lohnend ansehen. Den bisherigen Untersuchungen in der Literatur über den Zusammenhang zwischen Arbeitslohn und Arbeitsangebot folgend, führt diese individuelle Einschätzung über die Steuerbelastung dann zu der hier festgestellten Reduktion des Arbeitsangebots.³⁴

Alternativ kann das eingeschränkte Arbeitsangebot auch auf eine allgemeine Unzufriedenheit zurückzuführen sein, die bei den Teilnehmern durch die Konfrontation mit vielen Steuersätzen ausgelöst wird. Dadurch könnte die Arbeitsmotivation gesenkt werden, ohne überhaupt einen Steuersatz zu berechnen bzw. zu schätzen. Erste Implikationen für den Gesetzgeber sind hier bereits ersichtlich. Offensichtlich bewirkt nicht nur die Höhe der effektiven Steuerbelastung eine Anpassungsreaktion bei den Zensiten, sondern auch der Grad der Komplexität. Komplexe Steuertarifsysteme wirken dabei klar arbeitsangebotsreduzierend. Die Strategie, einzelne kleine Steuerbelastungen in bzw. hinter einem komplexen Steuersystem zu verstecken, um so das Belastungsgefühl der Zensiten zu senken, scheint vor diesem Hintergrund keine Erfolg versprechende Strategie zu sein, da die Komplexität selbst, zumindest in einem gewissen Rahmen, zu Steuerwiderstand führt. Dieser äußert sich in diesem Experiment durch eine Reduzierung des Arbeitsangebots.

Zusätzlich zu den Ergebnissen aus dem Experiment wurden die Teilnehmer im Anschluss noch gebeten, einen Fragebogen zu ihren Erfahrungen während des Experiments auszufüllen. Insbesondere wurden die Teilnehmer gefragt, ob sie die Vergütung fair fanden bzw. wie sie diese beurteilen, welchen Einfluss die Besteuerung und der entsprechende Nettolohn auf ihre Arbeitsbereitschaft hatten und warum sie das Experiment beendet haben. Außerdem wurden die Teilnehmer gebeten, den für sie geltenden Steuersatz noch einmal zu nennen, sofern sie diesen direkt ablesen konnten (Steuersystem A) oder diesen vor dem Experiment bereits berechnet hatten. Wenn sie den Steuersatz noch nicht vor Beginn des Experiments berechnet hatten, sollten sie dies jetzt nachholen. Die folgende Tabelle fasst die Mittelwerte der relevanten Fragen zusammen. Fragen hinsichtlich der Einschätzung der Tätigkeit, des Einflusses von Steuern und Nettolohn, der Fairness sowie der Steuerrechtskenntnisse wurde jeweils auf einer Skala von eins bis zehn abgefragt. Die entsprechenden Fragen finden sich im Anhang.

³⁴ Vgl. Fußnote 14.

Steuersystem:	A	B	C	D
Wie angenehm fanden Sie die Tätigkeit? (1 = sehr unangenehm; 10 = sehr angenehm)	4,96	4,21	4,54	3,75
Anteil der Teilnehmer, die ihren Steuersatz vor Beginn der Tätigkeit berechnet haben.	88%	83%	58%	38%
Nennen Sie uns jetzt bitte nochmals den Steuersatz bzw. berechnen sie diesen jetzt. ³⁵	58,33% (60%) (0,0816)	57,35% (60%) (0,0520)	59,06% (60%) (0,0804)	63,92% (64%) (0,1103)
Welchen Einfluss hatte die Steuer auf ihr Arbeitsangebot? (1 = gar keinen Einfluss; 10 = sehr großen Einfluss)	6,25	5,63	5,83	6,79
Nennen Sie uns jetzt bitte nochmals den Nettolohn bzw. berechnen Sie diesen jetzt.	12,78	14,41	13,71	13,32
Welchen Einfluss hatte der Nettolohn auf ihr Arbeitsangebot? (1 = gar keinen Einfluss; 10 = sehr großen Einfluss)	7,25	6,54	6,04	5,88
Wie fair fanden Sie die Vergütung? (1 = gar nicht fair; 10 = sehr fair)	4,58	4,00	3,92	4,54
Wie hoch schätzen Sie ihre eigenen Steuerrechtskenntnisse ein? (1 = keine Kenntnisse; 10 überdurchschnittliche Kenntnisse)	4,50	4,04	4,17	3,79

Tabelle 9 - Fragen zum Experiment

Die Tabelle zeigt, dass in Gruppe A der Steuersatz von 88% der Teilnehmer vor Beginn des Experiments berechnet bzw. beachtet wurde, wohingegen dieser Wert mit steigender Komplexität abnimmt.³⁶ In Gruppe D haben lediglich 38% der Teilnehmer vor dem Experiment ihren Steuersatz berechnet. Da bei Gruppe A der effektive Steuersatz explizit auf der Steuersystemkarte angegeben war, wäre hier ein Ergebnis von 100% zu erwarten. Diese Differenz ist wohl darauf zurückzu-

³⁵ Die Werte in Klammern geben die Mediane sowie die Standardabweichung bei dieser Frage an.

³⁶ Diese Frage zielte im Fragebogen darauf ab, ob der Steuersatz vor Aufnahme der Arbeitstätigkeit berechnet wurde. Da bei Gruppe A der effektive Steuersatz direkt auf der Steuersystemkarte angegeben wurde, sollte diese Frage bei der Gruppe Aufschluss darüber geben, ob dieser Steuersatz überhaupt von den Teilnehmern beachtet bzw. wahrgenommen wurde.

führen, dass 12% der Probanden die Frage verneint haben, da diese dem Steuersatz vor dem Experiment offensichtlich einfach keine Beachtung geschenkt haben. Auch bei der Berechnung des Steuersatzes liegt der Mittelwert in Gruppe A nicht genau bei 60%. Dies ist einem Probanden geschuldet, der trotz der konkreten Angabe des effektiven Steuersatzes auf der Steuersystemkarte einen Wert von 20% als effektive Steuerbelastung im Fragebogen angegeben hat. Deshalb wurden in die Tabelle ebenfalls die Mediane bei dieser Frage angegeben. Insbesondere in Gruppe D fällt auf, dass der Steuersatz im Durchschnitt überschätzt wurde. Hier liegen insgesamt 12 der 24 Gruppenteilnehmer mit ihrer Schätzung über den eigentlichen 60% Steuersatz. Drei Probanden haben hier einen Wert von 80% angegeben, was der Summe aller auf der Karte angegebenen Steuersätze entspricht und als Anzeichen für eine Vereinfachungsheuristik bei der Kalkulation des Steuersatzes gewertet werden kann. Auch die angegebene Standardabweichung zeigt, dass bei Gruppe D die angestellten Berechnungen bzw. Schätzungen weiter um den Mittelwert streuen, was ebenfalls als ein Zeichen für die Berechnungsschwierigkeiten gedeutet werden kann. Weiterhin fällt auf, dass nur in Gruppe D die Wichtigkeit des Nettolohns bei der Arbeitsangebotsentscheidung niedriger als die Wichtigkeit der Steuer selbst ist. Ursächlich dafür könnte die Tatsache sein, dass die Probanden hier den Nettolohn nicht oder nur in Ansätzen berechnen konnten, was wiederum dazu führt, dass alleine das Steuertarifsystem deren Arbeitsangebot, bspw. aufgrund der Konfrontation mit vier verschiedenen Steuerarten, gesenkt hat. Die Tätigkeit selbst scheint von den Teilnehmern in den ersten drei Steuertarifsystemgruppen auch im Mittel als angenehmer empfunden worden zu sein, was ebenfalls ein Indiz für die abschreckende Wirkung des komplexen Steuertarifsystems ist. Da die Probanden zufällig auf die Steuertarifsysteme verteilt wurden, scheint hier ebenfalls das Steuertarifsystem verantwortlich für die Zufriedenheit mit der konkreten Form der Arbeitsaufgabe zu sein.

Zur Bestätigung dieser Ergebnisse wurden wiederum parameterfreie Testverfahren durchgeführt. Damit konnten signifikante Unterschiede auf einem Signifikanzniveau von 5% bei der Frage nach der Höhe des Steuersatzes, bei der Frage, ob der Steuersatz oder der Nettolohn vorher berechnet wurden und bei der Frage, wie angenehm die Tätigkeit empfunden wurde, festgestellt werden. Bei den anderen Fragen gibt es keine signifikanten Unterschiede zwischen den einzelnen Steuertarifsystemgruppen.

Zum Schluss der Untersuchung sollen die Einflüsse des Steuertarifsystems sowie weiterer Größen auf das Arbeitsangebot mittels einer OLS Regression genauer quantifiziert werden. Eine Regression eignet sich dafür besonders, da so der Einfluss mehrerer erklärender Variablen gleichzeitig identi-

fiziert und analysiert werden kann. Als abhängige Variable wird dabei zum einen die Arbeitszeit und zum anderen die Arbeitsleistung (digitalisierte Bögen) verwendet. Die zu schätzende Regressionsgleichung ergibt sich somit als:

$$y_i = \beta_0 + \underbrace{\beta_1 \text{Dummy}_i^B + \beta_2 \text{Dummy}_i^C + \beta_3 \text{Dummy}_i^D}_{\text{Dummyvariablen für die Steuersystemgruppen}} + \underbrace{\beta_k \mathbf{X}_{i,k}}_{\text{Kontrollvariablen}} + \varepsilon_i,$$

wobei die Matrix \mathbf{X} die Kontrollvariablen: „Einfluss der Besteuerung“, „Zufriedenheit mit der Tätigkeit“, „Fairness der Vergütung“, „Steuerrechtskenntnisse“, „Alter“, „Geschlecht“, „Einkommensklassen“ sowie „Studium der Betriebswirtschaftslehre“ enthält.

Die folgenden Tabellen stellt die Ergebnisse der Regression dar. Dabei wurde die abhängige Variable jeweils in ihrer tatsächlichen Form wie auch in der um Produktivitätseffekte korrigierten Form verwendet.

Abhängige Variable	Tatsächliche Arbeitszeit		Korrigierte Arbeitszeit	
	Betakoeffizient	Standardfehler	Betakoeffizient	Standardfehler
Konstante	97,683***	34,105	91,201**	45,766
Steuersystem B	-11,279	12,576	-26,183	16,875
Steuersystem C	-29,760**	12,011	-30,748*	16,118
Steuersystem D	-48,646***	12,449	-57,774***	16,706
Einfluss der Besteuerung	-3,066*	1,733	-6,508***	2,325
Zufriedenheit mit der Tätigkeit	6,630***	2,408	11,328***	3,232
Fairness der Vergütung	1,356	2,326	4,356*	3,122
Steuerrechtskenntnisse	-0,370	2,254	-4,090	3,025
Alter	1,041	1,250	1,903	1,678
Geschlecht	-4,914	9,022	3,223	12,106
Einkommen 501 - 1.000 Euro	-16,464	10,519	-28,974**	14,115
Einkommen größer 1.000 Euro	-17,519	16,841	-29,064	22,599
Studium der Betriebswirtschaftslehre	3,357	11,765	10,948	15,788

***Signifikant auf dem 1% Niveau, **5% Niveau, *10% Niveau

Tabelle 10 - OLS Regression, Abhängige Variable: Arbeitszeit, Erweiterte Spezifikation

Da vermutet werden kann, dass die Kontrollvariablen „Einfluss der Besteuerung“, „Zufriedenheit mit der Tätigkeit“, „Fairness der Vergütung“ und „Steuerrechtskenntnisse“ in dem Modell endogen sind (da diese durch das Steuertarifsystem beeinflusst werden), wird diese Regression zur Überprüfung der Ergebnisse nochmals unter Ausschluss dieser Variablen durchgeführt.

Abhängige Variable	Tatsächliche Arbeitszeit		Korrigierte Arbeitszeit	
	Betakoeffizient	Standardfehler	Betakoeffizient	Standardfehler
Konstante	110,502***	31,202	119,023**	47,119
Steuersystem B	-16,021	13,323	-31,997	20,120
Steuersystem C	-32,945**	12,782	-34,178*	19,303
Steuersystem D	-58,833***	12,823	-72,756***	19,365
Alter	1,394	1,247	1,792	1,883
Geschlecht	-7,076	9,371	-2,996	14,151
Einkommen 501 - 1.000 Euro	-19,620*	11,178	-32,461*	16,880
Einkommen größer 1.000 Euro	-15,431	18,043	-23,445	27,247
Studium der Betriebswirtschaftslehre	3,700	11,395	5,806	17,208

***Signifikant auf dem 1% Niveau, **5% Niveau, *10% Niveau

Tabelle 11 - OLS Regression, Abhängige Variable: Arbeitszeit, Basisspezifikation

Die Regression zeigt, dass die Steuertarifsysteme C und D einen signifikanten Einfluss auf das Arbeitsangebot relativ zur Referenzgröße „Steuersystem A“ aufweisen. Die Zugehörigkeit zu einem dieser Steuertarifsysteme führt somit zu einer Senkung des Arbeitsangebotes. Dieses Ergebnis deckt sich mit den oben durchgeführten Gruppenvergleichen. Außerdem wird ersichtlich, dass Teilnehmer, die im Fragebogen einen höheren Einfluss der Besteuerung bei ihrer Entscheidungsfindung angegeben haben, etwas weniger gearbeitet haben als Teilnehmer, die der Besteuerung nur einen geringen Einfluss beigemessen haben. Dieses Ergebnis war insoweit zu erwarten und bestätigt den gefundenen Einfluss der Besteuerung auf die Arbeitsangebotsentscheidung weiter. Ebenfalls übt die Zufriedenheit mit der Art der Arbeitstätigkeit einen hoch signifikanten und positiven Einfluss auf die Arbeitszeit aus. Beim Einkommen wurde die Einkommensklasse „bis 500 Euro monatlich verfügbares Einkommen“ als Referenzkategorie gewählt. Nur für die korrigierte Arbeitszeit zeigt sich hier ein signifikanter negativer Einfluss bei der Einkommensklasse 501 – 1.000 Euro. Das Ergebnis für die andere Einkommensklasse ist dagegen nicht signifikant. Dieses Ergebnis ist auch insoweit schlüssig, als dass die Studierenden, die mehr Geld zur Verfügung haben, die Verdienstmöglichkeiten in dem Experiment vermutlich nicht als so hoch einschätzen, was wiederum deren Arbeitszeitangebot im Mittel senkte. Sofern diese ihr verfügbares Einkommen selbst verdienen, ist auch

von höheren Opportunitätskosten auszugehen. Für die tatsächliche Arbeitszeit ist dieser Einfluss zwar ebenfalls negativ, aber in der erweiterten Spezifikation nicht signifikant. Die Variablen „Alter“, „Geschlecht“ sowie „Studium der Betriebswirtschaftslehre“ üben keinen signifikanten Einfluss auf die Arbeitszeit aus. Insgesamt liefern beide Regressionen sowohl für die erweiterte Spezifikation wie auch für die Basisspezifikation ein ähnliches Ergebnis. Im Folgenden werden die gleichen Regressionen für die Messung der Arbeitsleistung (Bögen pro Minute) durchgeführt:

Abhängige Variable	Tatsächlich digitalisierte Bögen		Korrigierte Bögen	
	Betakoeffizient	Standardfehler	Betakoeffizient	Standardfehler
Konstante	62,596**	31,410	67,042***	23,407
Steuersystem B	-17,970	11,582	-7,741	8,631
Steuersystem C	-21,102*	11,062	-20,427**	8,244
Steuersystem D	-39,652***	11,466	-33,387***	8,544
Einfluss der Besteuerung.	-4,467***	1,596	-2,104*	1,189
Zufriedenheit mit der Tätigkeit	7,774***	2,218	4,550***	1,653
Fairness der Vergütung	2,990	2,143	,930	1,597
Steuerrechtskenntnisse	-2,807	2,076	-,254	1,547
Alter	1,306	1,151	,714	,858
Geschlecht	2,212	8,309	-3,373	6,192
Einkommen 501 - 1.000 Euro	-19,885**	9,688	-11,299	7,219
Einkommen größer 1.000 Euro	-19,947	15,510	-12,024	11,558
Studium der Betriebswirtschaftslehre	7,514	10,836	2,303	8,075

***Signifikant auf dem 1% Niveau, **5% Niveau, *10% Niveau

Tabelle 12 - OLS Regression, Abhängige Variable: Arbeitsleistung, Erweiterte Spezifikation

Abhängige Variable	Tatsächlich digitalisierte Bögen		Korrigierte Bögen	
	Betakoeffizient	Standardfehler	Betakoeffizient	Standardfehler
Konstante	81,688**	32,338	75,841***	21,414
Steuersystem B	-21,960	13,809	-10,995	9,144
Steuersystem C	-23,457**	13,248	-22,613**	8,773
Steuersystem D	-49,934***	13,290	-40,379***	8,801
Alter	1,230	1,292	,957	,856
Geschlecht	-2,055	9,712	-4,856	6,431
Einkommen 501 - 1.000 Euro	-22,279**	11,585	-13,465*	7,671
Einkommen größer 1.000 Euro	-16,091	18,700	-10,591	12,383
Studium der Betriebswirtschaftslehre	3,984	11,810	2,538	7,821

***Signifikant auf dem 1% Niveau, **5% Niveau, *10% Niveau

Tabelle 13 - OLS Regression, Abhängige Variable: Arbeitsleistung, Basisspezifikation

Auch für die Arbeitsleistung lässt sich ein ähnliches Ergebnis wie für die Arbeitszeit ableiten. Hier hat die Erhöhung der Komplexität des Steuertarifsystems ebenfalls einen signifikant negativen Einfluss auf das Arbeitsangebot. Dagegen wirkt sich auch hier die Zufriedenheit mit der Arbeitsaufgabe positiv auf das Arbeitsangebot aus. Der Einfluss der Besteuerung hat ebenfalls wieder einen leicht negativen Einfluss. Insgesamt decken sich diese Ergebnisse mit den vorherigen Regressionsergebnissen. Auch die zur Überprüfung durchgeführte Basisspezifikation zeigt wieder ein nahezu identisches Bild wie die erweiterte Spezifikation.

Es wird somit deutlich, dass nicht nur der effektive Steuersatz allein einen Einfluss auf das Verhalten der Zensiten ausübt. Auch der Komplexitätsgrad des Tarifsystems scheint einen maßgeblichen und quantifizierbaren Einfluss auszuüben, was hier am Beispiel der Arbeits-Freizeit-Entscheidung gezeigt werden konnte. Die Annahme von exakt kalkulierenden Steuerpflichtigen in theoretischen Modellen scheint vor diesem Hintergrund zumindest dann mehr als fraglich, wenn aus diesen Modellen auf das tatsächliche Verhalten in der Realität geschlossen werden soll. Die Teilnehmer haben sich offensichtlich Heuristiken und Vereinfachungsregeln bedient, welche ihr Verhalten maßgeblich determiniert haben. Auch die Richtung dieser Verhaltensbeeinflussung ist dabei deutlich erkennbar. So führt ein komplexes Steuertarifsystem zu einer Reduzierung des Arbeits-

angebotes und nicht etwa zu einer vereinfachten Orientierung an Vor-Steuer-Werten, welche dann zu einer Angebotsausweitung führen müssten.

Auch für die Steuerpolitik sind diese Ergebnisse von Bedeutung. Das Experiment hat gezeigt, dass es eben nicht nur auf den tatsächlichen effektiven Steuersatz ankommt. Vielmehr kann eine vereinfachte Darstellung des Steuertarifsystems dabei helfen, die richtigen Anreize bei den Zensiten zu setzen und komplexitätsbedingte Fehlanreize zu vermeiden. Diesem Punkt sollte in der Steuerreformdiskussion mehr Beachtung geschenkt werden. Den effektiven Steuersatz zu senken, dafür aber immer komplexere Regelungen zu implementieren, scheint vor diesem Hintergrund nicht das angemessene Vorgehen zu sein.

5. Zusammenfassung und Ausblick

Ziel des vorliegenden Beitrags war es, den Einfluss steuerlicher Tarifkomplexität auf die Arbeits-Freizeit-Entscheidung von Steuerpflichtigen zu untersuchen. Es wurde zu diesem Zweck ein Experiment durchgeführt. Dabei wurden reale monetäre Anreize gesetzt, um so das Verhalten der Teilnehmer unter realitätsnahen Bedingungen zu beobachten. Die Teilnehmer des Experiments wurden mit einer Arbeitsaufgabe konfrontiert, bei welcher diese die Möglichkeit hatten, einen Bruttolohn in Abhängigkeit von ihrer Arbeitsleistung zu erzielen. Weiterhin bekamen die Teilnehmer eines von vier verschiedenen komplexen Steuertarifsystemen zugeteilt, welches für die Besteuerung des von ihnen verdienten Arbeitslohnes zur Anwendung kam. Die Komplexität wurde dabei durch das Zusammenspiel verschiedener einzelner Steuersätze modelliert, welche additiv oder multiplikativ verknüpft werden mussten, um die effektive Gesamtsteuerbelastung zu berechnen. Der effektive Steuersatz war bei allen Varianten gleich, was dazu führte, dass auch der Nettolohn gleich war. Es konnte gezeigt werden, dass die Gruppe mit dem höchsten Steuerkomplexitätsgrad signifikant weniger arbeitete als die Gruppen mit niedrigeren Steuerkomplexitätsgraden, wobei sich auch dort Abstufungen beim Arbeitsangebot zeigen ließen. Das Arbeitsangebot wurde über die selbst gewählte Arbeitsdauer und über die jeweils erbrachte Arbeitsleistung modelliert. Auch mögliche exogene Produktivitätsunterschiede wurden bei der Analyse berücksichtigt. Im Ergebnis konnte gezeigt werden, dass Steuerpflichtige auf komplexe Tarifsysteme mit einer Reduktion ihres Arbeitsangebots reagieren. Tarifkomplexität führt demnach zu einer systematischen Abweichung vom theoretischen Modell eines exakt kalkulierenden Steuerpflichtigen. Dieses Ergebnis besitzt sowohl für die theoretische Forschung als auch für die steuerpolitische Praxis Relevanz und sollte bei der Erarbeitung von theoretischen Modellen und auch bei der Steuerreformdiskussion beachtet

werden. Vor allem die These, dass der Staat Abgaben durch die Komplexität des Steuersystems verschleiern und somit den Steuerwiderstand mindern kann, kann hiermit widerlegt werden.

Im Folgenden sollen noch die Grenzen dieser Untersuchung sowie zukünftiges Forschungspotenzial aufgezeigt werden. Zur Überprüfung der Ergebnisse sollte die Untersuchung mit einer größeren Anzahl an Personen durchgeführt werden. Auch die Auswahl der Probanden, welche hier eine homogene Gruppe darstellen, könnte noch hinsichtlich ihrer Repräsentativität ausgebaut werden. Nicht modelliert wurden Lerneffekte, die in der Praxis durch die wiederholte Beschäftigung mit Ertragsteuern auftreten könnten. Allerdings befindet sich weltweit die Einkommensbesteuerung seit jeher im Fluss, und bspw. in Deutschland erhalten neue Tarifnormen nahezu jährlich Einzug in das Einkommensteuerrecht. Deshalb scheinen Lerneffekte bei der Konzeption dieser Fragestellung keine wesentliche Rolle zu spielen. Wenn sich ein Steuerpflichtiger in der Realität die Mühe macht, seine Steuerbelastung exakt zu berechnen, um auf dieser Grundlage seine ökonomischen Entscheidungen zu treffen, ist es eher unwahrscheinlich, dass diese Rechnung im nächsten Jahr noch genauso aussehen wird. Es sei dabei nur an die Einführung verschiedener Schedulisierungsnormen in der letzten Zeit, wie bspw. der Abgeltungsteuer oder der Thesaurierungsbegünstigung, gedacht.

Anhang

Anleitung

Mit der Teilnahme an diesem Experiment haben Sie die Möglichkeit, Geld zu verdienen. Die Höhe des Verdienstes am Ende des Experiments richtet sich nach ihrer persönlichen Anstrengung. Bitte lesen Sie diese Anleitung aufmerksam durch. Bei Fragen wenden Sie sich bitte an die Experimentleitung.

Das Ziel dieses Experiments ist es, Informationen über Ihr persönliches Arbeitsangebot zu erhalten. Zu diesem Zweck werden Sie mit einer realen Arbeitssituation konfrontiert, bei welcher Sie Geld verdienen können. Zur Vergleichbarkeit der Daten verschiedener Teilnehmer handelt es sich bei der Arbeitsaufgabe um eine Tätigkeit, welche keine Vorkenntnisse oder speziellen Fähigkeiten erfordert. Wir möchten Sie darauf hinweisen, dass Gespräche mit anderen Teilnehmern, das Verlassen des Platzes sowie die Nutzung des PCs für private Zwecke während der Durchführung des Experiments nicht erlaubt sind. Nachdem Sie diese Anleitung gelesen haben, erhalten Sie mehrere Lösungsbögen einer Klausur, welche zu digitalisieren sind. Dazu finden Sie auf Ihrem Rechner eine vorgefertigte Exceldatei, in welche die Lösungsbögen zu übernehmen sind. Wie das Digitalisierungsverfahren konkret durchgeführt werden soll, wird Ihnen nach dem Lesen dieser Anleitung exemplarisch vom Experimentleiter vorgeführt.

*Abhängig von der Anzahl der fehlerfrei digitalisierten Bögen erhalten Sie am Ende des Experiments ihre Gehaltszahlung. Für jeden digitalisierten Bogen erhalten Sie einen **Bruttolohn in Höhe von 30 Cent**. Wenn Sie also beispielsweise durchschnittlich einen Bogen pro Minute digitalisieren, erhalten Sie ein Gehalt von 18 Euro brutto pro Stunde; bei durchschnittlich 1,5 Bögen pro Minute erhalten Sie ein Gehalt von 27 Euro brutto pro Stunde. Bitte beachten Sie: Von Ihrem Bruttolohn müssen Sie noch Steuerzahlungen leisten, nur der Rest (Nettolohn) wird Ihnen am Ende des Experiments bar ausgezahlt. Die Höhe der zu zahlenden Steuer bzw. den entsprechenden Steuersatz entnehmen Sie bitte der Steuertarifsystemkarte, welche Ihnen noch separat ausgehändigt wird. Lesen Sie auch diese bitte vor Beginn der Tätigkeit sorgfältig durch!*

Sie entscheiden selbst über Ihr Arbeitsangebot! Das bedeutet, dass Sie so lange und soviel arbeiten können wie Sie wollen. Sie können das Experiment also jederzeit beenden und erhalten im Anschluss daran Ihre Gehaltsauszahlung in Abhängigkeit von Ihrer erbrachten Arbeitsleistung. Bitte melden Sie sich beim Experimentleiter, wenn Sie das Experiment beenden wollen.

Viel Vergnügen!

Steuersystemkarten

Bei Aushändigung der Karten an die Probanden war jede Karte auf die gleiche Größe normiert.

A

- (1) Die Steuer auf das Einkommen beträgt 60%.

B

- (1) Die Steuer auf das Einkommen (Einkommensteuer) beträgt 40%.
- (2) Zusätzlich wird eine Zusatzsteuer auf das Einkommen in Höhe von 20% erhoben.

C

- (1) Die Steuer auf das Einkommen (Einkommensteuer) beträgt 40%.
- (2) Zusätzlich wird eine Zuschlagsteuer auf die Einkommensteuer (nicht auf das Einkommen selbst!) in Höhe von 20% erhoben.
- (3) Außerdem fällt eine Landessteuer in Höhe von 12% auf das Einkommen an.

D

- (1) Die Steuer auf das Einkommen (Einkommensteuer) beträgt 20%.
- (2) Des Weiteren wird eine Zuschlagsteuer auf die Einkommensteuer (nicht auf das Einkommen selbst!) in Höhe von 20% erhoben.
- (3) Zusätzlich fällt eine Landessteuer in Höhe von 20% auf das Einkommen an.
- (4) Weiterhin fällt eine Gemeindesteuer von ebenfalls 20% auf das Einkommen an.
- (5) Die Gemeindesteuer vermindert die Höhe des Einkommens, auf welches die Einkommensteuer erhoben wird (Abzugsfähigkeit der Gemeindesteuer bei der Berechnung der Einkommensteuer).

Fragen aus dem Fragebogen:

1. Wie fanden Sie die Tätigkeit, die Sie während des Experimentes ausüben sollten auf einer Skala von 1 = „sehr unangenehm“ bis 10 = „sehr angenehm“?

sehr unangenehm

1	2	3	4	5	6	7	8	9	10
---	---	---	---	---	---	---	---	---	----

sehr angenehm

2. Haben Sie vor Beginn der Arbeitstätigkeit Ihren Steuersatz ermittelt bzw. dieses versucht?

- Ja.
 Nein.

3. Nennen Sie uns bitte den für Sie geltenden effektiven (Gesamt-)Steuersatz jetzt noch einmal, bzw. versuchen Sie diesen bitte jetzt zu berechnen, wenn Sie dies noch nicht während des Experiments getan haben.

Steuersatz (in Prozent): _____%.

4. Welchen Einfluss hatte die erhobene Steuer auf Ihre Arbeitsbereitschaft auf einer Skala von 1 = „gar keinen Einfluss“ bis 10 = „sehr großen Einfluss“.

gar keinen Einfluss

1	2	3	4	5	6	7	8	9	10
---	---	---	---	---	---	---	---	---	----

sehr großen Einfluss

5. Haben Sie vor Beginn des Experiments Ihren Nettolohn pro Bogen ermittelt?

- Ja.
 Nein.

6. Wie wichtig war für Sie die Höhe des Nettolohns bei der Entscheidung zu arbeiten auf einer Skala von 1 = „ganz unwichtig“ bis 10 = „sehr wichtig“.

ganz unwichtig

1	2	3	4	5	6	7	8	9	10
---	---	---	---	---	---	---	---	---	----

sehr wichtig

7. Wie fair fanden Sie die Vergütung für die von Ihnen geforderte Tätigkeit auf einer Skala von 1 = „gar nicht fair“ bis 10 = „sehr fair“.

gar nicht fair

1	2	3	4	5	6	7	8	9	10
---	---	---	---	---	---	---	---	---	----

sehr fair

8. Wie schätzen Sie Ihre eigenen Steuerrechtskenntnisse auf einer Skala von 1 = „keine Kenntnisse“ bis 10 = „überdurchschnittliche Kenntnisse“ ein?

Keine Kenntnisse

1	2	3	4	5	6	7	8	9	10
---	---	---	---	---	---	---	---	---	----

Überdurchschnittliche Kenntnisse

9. Wenn Sie gerade studieren, nennen Sie uns bitte Ihren Studiengang sowie ihr aktuelles Fachsemester:

- Studiengang:_____.
- Fachsemester:_____.

10. Haben Sie in Ihrem Studium bereits Veranstaltungen im Vertiefungsgebiet „Steuerlehre“ besucht?

- Ja.
- Nein.

11. Geschlecht:

- männlich
- weiblich

12. Alter:

_____ Jahre

13. Wie hoch ist ungefähr Ihr monatliches zur Verfügung stehendes Einkommen?

- < 500 €
- 501 € – 1.000 €
- 1.001 € – 1.500 €
- 1.501 € – 2.000 €
- > 2.000 €

Literaturverzeichnis

- Alm, J. (2010), Testing Behavioral Public Economics Theories in the Laboratory. *National Tax Journal*, 63 (4), 635-658.
- Alm, J., B. Jackson, M. McKee (1992), Institutional Uncertainty and Taxpayer Compliance. *American Economic Review*, 82 (4), 1018-1026.
- Atkinson, A. B., J. E. Stiglitz, (1976), The design of tax structure: direct versus indirect taxation. *Journal of Public Economics*, 6 (1), 55-75.
- Bach, S. (2011), Kirchhof macht sich das Steuerrecht einfach. Kommentar, *DIW Wochenbericht*, 28/2011, 13. Juli 2011.
- Bartolome, C. A. de (1995), Which tax rate do people use: Average or marginal? *Journal of Public Economics*, 56, 79-96.
- Beck, P. J., J. S. Davis, W.-O. Jung (1991), Experimental evidence on taxpayer reporting under uncertainty. *Accounting Review*, 66 (3), 535-558.
- Bizer, K. (2004), Das Verhältnis von Komplexität und Steuerehrlichkeit: experimentelle Evidenz, in: Bizer, K., A. Falk, J. Lange (Hrsg.), *Am Staat vorbei - Partizipation, Fairness und Transparenz contra Steuerhinterziehung*. Berlin, 59-90.
- Bizer, K. (2008), *Steuervereinfachung und Steuerhinterziehung – eine experimentelle Analyse zur Begründung von Steuereinfachheit*. Duncker & Humblot, Berlin.
- Blaufus, K., J. Bob., J. Hundsdoerfer, D. Kiesewetter, J. Weimann (2009), It's All About Tax Rates – An Empirical Study of Tax Perception. *arqus-Diskussionsbeitrag*, Nr. 106.
- Blaufus, K., R. Ortlieb (2009), Is Simple Better? A Conjoint Analysis of the Effects of Tax Complexity on Employee Preferences Concerning Company Pension Plans. *Schmalenbach Business Review*, 61, 60-83.
- Blumkin, T., B. J. Ruffle, Y. Ganun (2008), Are Income and Consumption Taxes Ever Really Equivalent? Evidence from a Real-Effort Experiment with Real Goods. *Working Paper*.

- Boylan, S. J., P. J. Frischmann (2006), Experimental Evidence on the Role of Tax Complexity in Investment Decision. *Journal of the American Taxation Association*, 28, 69-88.
- Break, G. F. (1957), Income Taxes and Incentives to Work: An Empirical Study. *The American Economic Review*, 47 (5), 529-549.
- Chetty, R., A. Looney, K. Kroft (2009), Salience and taxation: theory and evidence. *The American Economic Review*, 99, 1145-1177.
- Chetty, R., E. Saez (2009), Teaching the tax code: Earning responses to an Experiment with EITC Recipients. *NBER Working Paper*, Nr. 14836.
- Congdon, W. J., J. R. Kling, S. Mullainathan (2009), Behavioral Economics and Tax Policy. *National Tax Journal*, 62, 375-386.
- Epley, N., T. Gilovich (2006), The Anchoring-and-Adjustment Heuristic: Why the Adjustments Are Insufficient. *Psychological Science*, 17 (4), 311-318.
- FDP (2010), Für faire Finanzbeziehungen zwischen Bürger und Staat. Ob als Steuerzahler oder als Transferempfänger – Leistung muss sich immer lohnen. Beschluss des 61. Ordentlichen Bundesparteitags der FDP.
- Fochmann, M., J. Weimann, K. Blaufus, J. Hundsdoerfer, D. Kiesewetter (2010), Gross-wage illusion in a real effort experiment. *FEMM Working Paper*, No. 09.
- Fujii, E. T., C. B. Hawley (1988), On the Accuracy of Tax Perceptions. *Review of Economics & Statistics*, 70, 344-347.
- Gamage, D., A Hayashi, B. K. Nakamura (2010), Experimental Evidence of Tax Framing Effects on the Work/Leisure Decision. *Working Paper*.
- Gensemer, B. L., J. A. Lean, W. B. Neenan (1965), Awareness of Marginal Income Tax Rates Among High-Income Taxpayers. *National Tax Journal*, 18, 258-267.
- Gigerenzer, G., W. Gaissmaier (2011), Heuristic decision making. *Annual Review of Psychology*, 62, 451-482.

- Harberger, A. C. (1964), The Measurement of Waste. *American Economic Review*, 54 (3), 58-76.
- Hausman, J. (1981), Labor Supply, in: Aaron, H., J. Pechman (Hrsg.), How Taxes Affect Economic Behavior. Washington DC: The Brookings Institution, 27-72.
- Hechtner, F. (2010), Eine theoretische und empirische Studie über Einkommensteuertarife aus Sicht der Wirtschaftswissenschaft – Progressionswirkungen der synthetischen Einkommensteuer, Schedulenbesteuerung und Vermeidung von Doppelbesteuerungen, Berlin 2010.
- Hechtner, F., J. Hundsdoerfer (2009), Schedulenbesteuerung von Kapitaleinkünften mit der Abgeltungsteuer: Belastungswirkungen und neue Problemfelder. *Steuer und Wirtschaft*, 23-41.
- Hechtner, F., J. Hundsdoerfer, C. Sielaff (2011), Progressionseffekte und Varianten zur optimalen Steuerplanung bei der Thesaurierungsbegünstigung – Eine Abweichungsanalyse. *Zeitschrift für betriebswirtschaftliche Forschung*, 214-239.
- Hicks, J. R. (1935), The Theory of Wages. MacMillan, London.
- Houben, H., R. Maiterth (2008), Optimale Nutzung und Wirkungen von § 34a EStG. *Steuer und Wirtschaft*, 228-237.
- Hundsdoerfer, J., C. Sichtmann (2007), Zur Gewichtung von Steuern in unternehmerischen Investitionskalkülen: Eine Conjoint-Analyse des Entscheidungsverhaltens selbständiger Ärzte. *Betriebswirtschaftliche Forschung und Praxis*, 604-621.
- Hundsdoerfer, J., C. Sielaff, K. Blaufus, D. Kiesewetter, J. Weimann (2010), The Name Game for Contributions – Influence of Labeling and Earmarking on the Perceived Tax Burden. *FACTS Discussion Paper*, FU Berlin, 2010/12.
- Keane, M. (2010), Labor Supply and Taxes: A Survey. *UTS Working Paper*, 160.
- Keller, T. (1968), Merkliche und unmerkliche Steuern. *Finanzarchiv*, 100-120.
- Killingsworth, M. (1983), Labor Supply. Cambridge University Press, Cambridge.

- Kimmel, J., T. J. Kniesner (1998), New Evidence on Labor Supply: Employment vs. Hours Elasticities by Sex and Marital Status. *Journal of Monetary Economics*, 42, 289-301.
- Kirchhof, P. (2011), Bundessteuergesetzbuch – Ein Reformentwurf zur Erneuerung des Steuerrechts. Schriftenreihe des Instituts für Finanz- und Steuerrecht: Forschungsstelle Bundessteuergesetzbuch, C. F. Müller, Heidelberg.
- Koalitionsvertrag zwischen CDU, CSU und FDP (2009), Wachstum. Bildung. Zusammenhalt, 17. Legislaturperiode.
- Lewis, A. (1978), Perceptions of Tax Rates. *British Tax Review*, 6, 358-366.
- Maiterth, R. (2003), Beurteilung unterschiedlicher Reformvorschläge für die Gewerbesteuer aus betriebswirtschaftlicher Sicht. *Die Betriebswirtschaft*, 63, 479-496.
- Mann, H., D. Whitney (1947), On a test of whether one of two random variables is stochastically larger than the other. *Annals of Mathematical Statistics*, 18, 50-60.
- McCaffery, E. J., J. Baron (2003), The Humpty Dumpty blues: Disaggregation bias in the evaluation of tax systems. *Organizational Behavior and Human Decision Processes*, 91 (2), 230-242.
- Mirrlees, J. A. (1971), An Exploration in the Theory of Optimum Income Taxation. *The Review of Economic Studies*, 38 (2), 175–208.
- Neumark, F. (1970), Grundsätze gerechter und ökonomisch rationaler Steuerpolitik. Mohr/Siebeck, Tübingen.
- Plott, C. R. (1987), Dimensions of Parallelism: Some Policy Applications of Experimental Methods, in: Roth, A., E. Roth (Hrsg.), *Laboratory Experimentation in Economics: Six Points of View*. Cambridge University Press, New York, NY, 193-229.
- Ramsey, F. P. (1927), A Contribution to the Theory of Taxation. *Economic Journal*, 37 (145), 47-61.
- Rupert, T. J., A. M. Wright (1998), The Use of Marginal Tax Rates in Decision Making: The Impact of Tax Rate Visibility. *Journal of the American Taxation Association*, 20, 83-99.

- Rupert, T. J., C. M. Fischer (1995), An Empirical Investigation of Taxpayer Awareness of Marginal Tax Rates. *Journal of the American Taxation Association*, 17, 36-59.
- Schmölders, G. (1960), Das Irrationale in der öffentlichen Finanzwirtschaft. Rowohlt, Hamburg.
- Sillamaa, M.-A. (1999), Taxpayer behavior in response to taxation: comment and new experimental evidence. *Journal of Accounting and Public Policy*, 18, 165-177.
- Swenson, C. W. (1988), Taxpayers behavior in response to taxation: An experimental analysis. *Journal of Accounting and Public Policy*, 7, 1-28.
- Tipke, K., J. Lang (2010), Steuerrechtsordnung. 20. Aufl., Verlag Dr. Otto Schmidt, Köln.
- Tversky, A., D. Kahneman (1974), Judgment under Uncertainty: Heuristics and Biases. *Science*, 185 (4157), 1124-1131.
- Wagner, F. (1984), Grundfragen und Entwicklungstendenzen der betriebswirtschaftlichen Steuerplanung. *Betriebswirtschaftliche Forschung und Praxis*, 201-221.
- Wagner, F. (1999), Besteuerung, in: Vahlens Kompendium der Betriebswirtschaftslehre, Bd. 2, 4. Aufl., München, 439-504.
- Wagner, F. (2011), Der große Wurf ins Aus, Das Steuerkonzept Kirchhofs ist unklar, widersprüchlich und teuer. *Süddeutsche Zeitung*, 67. Jg., 158, 12. Juli 2011.
- Wilcoxon, F. (1945), Individual Comparisons by Ranking Methods. *Biometrics Bulletin*, 1, 80-83.
- Ziliak, J. P., T. J. Kniesner (1999), Estimating Life-Cycle Labor Supply Tax Effects. *Journal of Political Economy*, 107, 326-359.

Progressionseffekte und Varianten zur optimalen Steuerplanung bei der Thesaurierungsbegünstigung – Eine Abweichungsanalyse

Frank Hechtner, Jochen Hundsdoerfer, Christian Sielaff

Aus Urheberrechtsgründen ist dieser Beitrag in der Online-Version nicht enthalten. Der Aufsatz ist in *Schmalenbachs Zeitschrift für betriebswirtschaftliche Forschung* 2011 (5), S. 214-239 erschienen, <http://www.fachverlag.de/zfbf/>.

Charitable Giving in the German Welfare State: Fiscal Incentives and Crowding Out

Timm Bönke, Nima Massarrat-Mashhadi, Christian Sielaff

Aus Urheberrechtsgründen ist dieser Beitrag in der Online-Version nicht enthalten. Der Aufsatz ist in *Public Choice* 2011 (online first, DOI: 10.1007/s11127-011-9806-y) erschienen, <http://www.springer.com/>.